

Geometrisation of Statistical Mechanics

Dorje C. Brody* and Lane P. Hughston†

**Department of Applied Mathematics and Theoretical Physics,
University of Cambridge, Silver Street, Cambridge CB3 9EW U.K.*

*† Merrill Lynch International, 25 Ropemaker Street, London EC2Y 9LY U.K.
and King's College London, The Strand, London WC2R 2LS, U.K.*

(February 7, 2008)

Abstract. Classical and quantum statistical mechanics are cast here in the language of projective geometry to provide a unified geometrical framework for statistical physics. After reviewing the Hilbert space formulation of classical statistical thermodynamics, we introduce projective geometry as a basis for analysing probabilistic aspects of statistical physics. In particular, the specification of a canonical polarity on RP^n induces a Riemannian metric on the state space of statistical mechanics. In the case of the canonical ensemble, we show that equilibrium thermal states are determined by the Hamiltonian gradient flow with respect to this metric. This flow is concisely characterised by the fact that it induces a projective automorphism on the state manifold. The measurement problem for thermal systems is studied by the introduction of the concept of a random state. The general methodology is then extended to include the quantum mechanical dynamics of equilibrium thermal states. In this case the relevant state space is complex projective space, here regarded as a real manifold endowed with the natural Fubini-Study metric. A distinguishing feature of quantum thermal dynamics is the inherent multiplicity of thermal trajectories in the state space, associated with the nonuniqueness of the infinite temperature state. We are then led to formulate a geometric characterisation of the standard KMS-relation often considered in the context of C^* -algebras. The example of a quantum spin one-half particle in heat bath is studied in detail.

Keywords: Hilbert space geometry, Projective geometry, Equilibrium statistical mechanics, Quantum dynamics

I. INTRODUCTION

One of the most fascinating advances in the application of modern differential geometry is its use in statistical physics, including quantum and statistical mechanics. The purpose of this paper is to develop a unified geometrical framework that allows for a natural characterisation of both of these aspects of statistical physics.

In quantum mechanics, one typically works with square-integrable wave functions, i.e., elements of a complex Hilbert space \mathcal{H} . This space possesses natural geometrical structures induced by its norm. However, in order to seek a compelling axiomatic formulation of quantum mechanics, it may be reasonable to work with a space of more direct physical relevance [1,2]. This is not the Hilbert space \mathcal{H} itself, but rather the manifold Σ of “instantaneous pure states” [3], which has the structure of a complex projective space CP^n , possibly infinite dimensional, enriched with a Hermitian correspondence, i.e., a complex conjugation

operation that maps points to hyperplanes in CP^n , and vice-versa. Equivalently, we think of CP^n as being endowed with a natural Riemannian metric, the Fubini-Study metric.

The space Σ is, in fact, the quantum analogue of the classical phase space of mechanical systems. Hence, one can interpret the Schrödinger equation as Hamilton's equations on CP^n , and the equation of motion for a general density matrix can be identified with the Liouville equation [4]. The advantage of working with the manifold Σ , rather than the Hilbert space of state vectors, above all, is that it can readily accommodate generalisations of quantum mechanics [5], including nonlinear relativistic models. Furthermore, the structure of Σ allows for a natural probabilistic interpretation even if the standard linear quantum theory is modified.

As we discuss elsewhere [6], the statistical aspects of quantum measurement can be greatly clarified if we shift our view slightly, and regard the Hilbert space \mathcal{H} of quantum mechanics not as a complex Hilbert space, but rather a real Hilbert space endowed with a real metric and a compatible complex structure. This would appear to be simply a change in formalism while keeping the same underlying physical structure. Indeed this is so, but once quantum theory is formulated this way its relation to other aspects of statistical physics becomes much more apparent.

Statistical mechanics, in particular, can also be formulated concisely [7] in terms of the geometry of a real Hilbert space \mathcal{H} . This can be seen by taking the square root of the Gibbs density function, which maps the space of probability distributions to vectors in a convex cone \mathcal{H}_+ in \mathcal{H} . In this way, the various probabilistic and statistical operations of statistical mechanics can be given a transparent geometric meaning in \mathcal{H} [8,9].

However, it can be argued that even at the classical level of statistical mechanics the 'true' state space is obtained by identifying all the pure states along the given ray through the origin of \mathcal{H} . In this case, the space obtained is essentially the real projective space RP^n . This is the view we take here, and we shall study properties of thermal states that become apparent only when the theory is developed in a fully geometric context.

The present paper is organised as follows. In Section 2, we review the basics of the Hilbert space formulation of statistical mechanics. Since this formulation is perhaps not very widely appreciated, we can regard this section as an extended introduction which then paves the way to the approach in terms of projective geometry presented later. We begin with a brief review of statistical geometry, including the theory of the Fisher-Rao metric on the parameter space of a family of probability distributions. The Gibbs distribution when viewed in this way can be seen as a curve in Hilbert space, parameterised by the inverse temperature, and there is a striking formal resemblance to the Schrödinger equation, even though here we are working at a strictly classical level.

A measurement theory for thermal states is developed by analogy with the standard density matrix theory used in quantum mechanics. We are then led to a set of uncertainty relations for the measurements of thermodynamic conjugate variables such as energy and inverse temperature. We also introduce an alternative approach to the measurement theory that is not based upon the density matrix description. Our approach, based on the introduction of random states, extends naturally also to quantum mechanics, where it can be seen to be more appealing in a probabilistic context than standard treatments, and indeed reduces to the conventional density matrix approach in special cases.

In Section 3, we introduce a projective geometric framework for the probabilistic op-

erations involved in the representation of the canonical thermal state associated with the standard Gibbs measure. Thermal states are shown to lie on a trajectory in the real projective space RP^n , which is endowed with the natural ‘spherical’ metric. In this connection we find it convenient to develop a number of useful differential geometric results characterising projective transformations on the state space RP^n . We find that the equilibrium thermal trajectories, which are shown to be given by a Hamiltonian gradient flow, generate projective automorphisms of the state manifold.

In Section 4, we then synthesise the approaches outlined in Sections 2 and 3, and consider the inter-relationship of the classical thermal state space RP^n and the quantum phase space CP^n , to study the quantum mechanical dynamics of equilibrium thermal states. First we examine the quantum state space from the viewpoint of complex algebraic geometry, which shows that this space is endowed with a natural Riemannian geometry given by the Fubini-Study metric, along with a natural symplectic structure. For thermal physics it is instructive to look at quantum mechanics from an entirely ‘real’ point of view as well, and this approach is developed in Section 4.B.

Our formulation is then compared to the standard KMS-construction [10] for equilibrium states. In particular, once we pass to the mixed state description we recover the KMS-state. However, our quantum mechanical pure thermal state, which does not obey the KMS-condition, can be viewed as a more fundamental construction. In Section 4.D we develop a theory of the quantum mechanical microcanonical ensemble, formulated entirely in terms of the quantum phase space geometry. This is set up in such a way as to admit generalisations to nonlinear quantum theories. Finally, we study more explicitly the case of a quantum mechanical spin one-half particle in heat bath.

II. STATISTICAL STATES IN HILBERT SPACE

A. Hilbert space geometry

Let us begin by demonstrating how classical statistical mechanics can be formulated in an appealing way by the use of a geometrical formalism appropriate for Hilbert space. Consider a real Hilbert space \mathcal{H} equipped with an inner product g_{ab} . A probability density function $p(x)$ can be mapped into \mathcal{H} by taking the square-root $\psi(x) = (p(x))^{1/2}$, which is denoted by a vector ψ^a in \mathcal{H} . The normalisation condition $\int (\psi(x))^2 dx = 1$ is written $g_{ab}\psi^a\psi^b = 1$, indicating that ψ^a lies on the unit sphere \mathcal{S} in \mathcal{H} . Since a probability density function is nonnegative, the image of the map $f : p(x) \rightarrow \psi(x)$ is the intersection $\mathcal{S}_+ = \mathcal{S} \cap \mathcal{H}_+$ of \mathcal{S} with the convex cone \mathcal{H}_+ formed by the totality of quadratically integrable nonnegative functions. If we consider the space of all probability distributions as a metric space relative to the Hellinger distance [11], then f is an isometric embedding in \mathcal{H} . We call ψ^a the state vector of the corresponding probability density $p(x)$.

A typical random variable is represented on \mathcal{H} by a symmetric tensor X_{ab} , whose expectation in a normalised state ψ^a is given by

$$E_\psi[X] = X_{ab}\psi^a\psi^b. \quad (1)$$

Similarly, the expectation of its square is $X_{ac}X_{b\bar{c}}\psi^a\psi^b$. The variance of X_{ab} in the state ψ^a

is therefore $\text{Var}_\psi[X] = \tilde{X}_{ac}\tilde{X}_b^c\psi^a\psi^b$, where $\tilde{X}_{ab} = X_{ab} - g_{ab}E_\psi[X]$ represents the deviation of X_{ab} from its mean in the state ψ^a .

We consider now the unit sphere \mathcal{S} in \mathcal{H} , and within this sphere a submanifold \mathcal{M} given parametrically by $\psi^a(\theta)$, where θ^i ($i = 1, \dots, r$) are local parameters. In particular, later on we have in mind the case where the parameter space spanned by θ^i represents the space of coupling constants in statistical mechanics associated with the given physical system. In the case of the canonical Gibbs measure there is a single such parameter, corresponding to the inverse temperature variable $\beta = 1/k_B T$. We write ∂_i for $\partial/\partial\theta^i$. Then, in local coordinates, there is a natural Riemannian metric \mathcal{G}_{ij} on the parameter space \mathcal{M} , induced by g_{ab} , given by $\mathcal{G}_{ij} = g_{ab}\partial_i\psi^a\partial_j\psi^b$. This can be seen as follows. First, note that the squared distance between the endpoints of two vectors ψ^a and η^a in \mathcal{H} is $g_{ab}(\psi^a - \eta^a)(\psi^b - \eta^b)$. If both endpoints lie on \mathcal{M} , and η^a is obtained by infinitesimally displacing ψ^a in \mathcal{M} , i.e., $\eta^a = \psi^a + \partial_i\psi^a d\theta^i$, then the separation ds between the two endpoints on \mathcal{M} is $ds^2 = \mathcal{G}_{ij}d\theta^i d\theta^j$, where \mathcal{G}_{ij} is given as above.

The metric \mathcal{G}_{ij} is, up to a conventional, irrelevant factor of four, the so-called Fisher-Rao metric on the space of the given family of distributions. The Fisher-Rao metric is usually defined in terms of a rather complicated expression involving the covariance matrix of the gradient of the log-likelihood function; but here we have a simple, transparent geometrical construction. The Fisher-Rao metric is important since it provides a geometrical basis for the key links between the statistical and physical aspects of the systems under consideration.

B. Thermal trajectories

Now suppose we consider the canonical ensemble of classical statistical mechanics, in the case for which the system is characterised by a configuration space and an assignment of an energy value for each configuration. The parametrised family of probability distributions then takes the form of the Gibbs measure

$$p(H, \beta) = q(x) \exp[-\beta H(x) - W(\beta)] , \quad (2)$$

where the variable x ranges over the configuration space, $H(x)$ represents the energy, $W(\beta)$ is a normalisation factor, and $q(x)$ determines the distribution at $\beta = 0$, where β is the inverse temperature parameter.

We now formulate a Hilbert space characterisation of this distribution. Taking the square-root of $p(H, \beta)$, we find that the state vector $\psi^a(\beta)$ in \mathcal{H} corresponding to the Gibbs distribution (2) satisfies the differential equation

$$\frac{\partial\psi^a}{\partial\beta} = -\frac{1}{2}\tilde{H}_b^a\psi^b , \quad (3)$$

where $\tilde{H}_{ab} = H_{ab} - g_{ab}E_\psi[H]$. Here the operator H_{ab} in the Hilbert space \mathcal{H} corresponds to the specified Hamiltonian function $H(x)$ appearing in (2). The solution of this equation can be represented as follows:

$$\psi^a(\beta) = \exp\left[-\frac{1}{2}(\beta H_b^a + \tilde{W}(\beta)\delta_b^a)\right] q^b , \quad (4)$$

where $\tilde{W}(\beta) = W(\beta) - W(0)$ and $q^a = \psi^a(0)$ is the prescribed distribution at $\beta = 0$.

Since $\psi^a(\beta)$ respects the normalisation $g_{ab}\psi^a\psi^b = 1$, for each value of the temperature β we find a point on \mathcal{M} in \mathcal{S}_+ . To be more specific, the thermal system can be described as follows. Consider a unit sphere \mathcal{S} in \mathcal{H} , whose axes label the configurations of the system, each of which has a definite energy. We let u_k^a denote an orthonormal basis in \mathcal{H} . Here, the index k labels all the points in the phase space of the given statistical system. In other words, for each point in phase space we have a corresponding basis vector u_k^a in \mathcal{H} for some value of k . With this choice of basis, a classical thermal state $\psi^a(\beta)$ can be expressed as a superposition

$$\psi^a(\beta) = e^{-\frac{1}{2}W(\beta)} \sum_k e^{-\frac{1}{2}\beta E_k} u_k^a, \quad (5)$$

where E_k is the energy for k -th configuration, and thus $\exp[W(\beta)] = \sum_k \exp(-\beta E_k)$ is the partition function. We note that the states u_k^a are, in fact, the energy eigenstates of the system, with eigenvalues E_k . That is to say, $H_b^a u_k^b = E_k u_k^a$. The index k in these formulae is formal in the sense that the summation may, if appropriate, be replaced by an integration. By comparing equations (4) and (5), we find that the initial ($\beta = 0$) thermal state q^a is

$$q^a = e^{-\frac{1}{2}W(0)} \sum_k u_k^a, \quad (6)$$

which corresponds to the centre point in \mathcal{S}_+ . This relation reflects the fact that all configurations are equally probable likely at infinite temperature.

Viewed as a function of β , the state trajectory $\psi^a(\beta)$ thus commences at the centre point q^a , and follows a curve on \mathcal{S} generated by the Hamiltonian H_{ab} according to (3). It is interesting to note that the curvature of this trajectory, given by

$$K_\psi(\beta) = \frac{\langle \tilde{H}^4 \rangle}{\langle \tilde{H}^2 \rangle^2} - \frac{\langle \tilde{H}^3 \rangle^2}{\langle \tilde{H}^2 \rangle^3} - 1, \quad (7)$$

arises naturally in a physical characterisation of the accuracy bounds for temperature measurements. This point is pursued further in Section 2.C below, and in [7]. In equation (7) the expression $\langle \tilde{H}^n \rangle$ denotes the n -th central moment of the observable H_{ab} . Here the curvature of the curve $\psi^a(\beta)$, which is necessarily positive, is the square of the ‘acceleration’ vector along the state trajectory $\psi^a(\beta)$, normalised by the square of the velocity vector.

C. Measurement for thermal states

Given the thermal trajectory $\psi^a(\beta)$ above, we propose, in the first instance, to consider measurement and estimation by analogy with the von Neumann approach in quantum mechanics. According to this scheme the *general* state of a thermodynamic system is represented by a ‘density matrix’ ρ^{ab} which in the present context should be understood to be a symmetric, semidefinite matrix with trace unity; that is to say, $\rho^{ab}\xi_a\xi_b \geq 0$ for any covector ξ_a , and $\rho^{ab}g_{ab} = 1$. Then, for example, we can write

$$E_\rho[X] = X_{ab}\rho^{ab} \quad (8)$$

for the expectation of a random variable X_{ab} in the state ρ^{ab} , and

$$\text{Var}_\rho[X] = X_{ab}X_c^b\rho^{ac} - (X_{ab}\rho^{ab})^2 \quad (9)$$

for the variance of X_{ab} in that state. It should be evident that in the case of a pure state, for which ρ^{ab} is of the form $\rho^{ab} = \psi^a\psi^b$ for some state vector ψ^a , these formulae reduce to the expressions considered in Section 2.A.

In particular, let us consider measurements made on a pure equilibrium state $\psi^a(\beta)$. Such measurements are characterised by projecting the prescribed state onto the ray in the Hilbert space corresponding to a specified point in the phase space. Hence, the probability of observing the k -th state, when the system is in the pure state ψ^a , is given by the corresponding Boltzmann weight

$$p_k = (g_{ab}\psi^a u_k^b)^2 = e^{-\beta E_k - W(\beta)} . \quad (10)$$

In terms of the density matrix description, the state before measurement is given by the degenerate pure state matrix $\rho^{ab} = \psi^a\psi^b$, for which the thermal development is

$$\frac{d\rho^{ab}}{d\beta} = -\frac{1}{2}(\tilde{H}_c^a\rho^{bc} + \tilde{H}_c^b\rho^{ac}) , \quad (11)$$

or equivalently $d\rho/d\beta = -\{\tilde{H}, \rho\}$, where $\{A, B\}$ denotes the symmetric product between the operators A and B . After a measurement, ρ^{ab} takes the form of a mixed state, characterised by a nondegenerate diagonal density operator for which the diagonal elements are the Boltzmann weights p_k . In this state vector reduction picture, the von Neumann entropy $-\text{Tr}[\rho \ln \rho]$ changes from 0 to its maximum value $S = \beta\langle H \rangle + W_\beta$, which can be viewed as the quantity of information gained from the observation.

More generally, suppose we consider the measurement of an arbitrary observable X_{ab} in the state $\psi^a(\beta)$ in the situation when the spectrum of X_{ab} admits a continuous component. In this case, we consider the spectral measure associated with the random variable X_{ab} . Then, the probability density for the measurement outcome x is given by the expectation $p(x, \beta) = \Pi_{ab}(X, x)\psi^a\psi^b$ of the projection operator

$$\Pi_b^a(X, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp[i\lambda(X_b^a - x\delta_b^a)] d\lambda . \quad (12)$$

In other words, we assign a projection-valued measure $\Pi_X(x)$ on the real line associated with each symmetric operator X , so that for a given unit vector ψ^a , the mapping $x \in \mathbf{R} \mapsto E_\psi[\Pi_X(x)]$ is a probability measure. This measure determines the distribution of values obtained when the observable X is measured while the system is in the state ψ .

For a more refined view of the measurement problem we need to take into account some ideas from statistical estimation theory. Suppose that we want to make a measurement or series of measurements to estimate the value of the parameter characterising a given thermal equilibrium state. In this situation the observable we measure is called an ‘estimator’ for the given parameter. We are interested in the case for which the estimator is unbiased in the sense that its expectation gives the value of the required parameter. To be specific, we consider the case when we estimate the value of the temperature. Let B_{ab} be an unbiased estimator for β , so that along the trajectory $\psi^a(\beta)$ we have:

$$\frac{B_{ab}\psi^a\psi^b}{g_{cd}\psi^c\psi^d} = \beta . \quad (13)$$

As a consequence of this relation and the thermal state equation (3), we observe that the inverse temperature estimator B and the Hamiltonian H satisfy the ‘weak’ anticommutation relation $E_\psi[\{B, \tilde{H}\}] = -1$ along the state trajectory ψ . In statistical terms, this implies that these conjugate variables satisfy the covariance relation $E_\psi[BH] - E_\psi[B]E_\psi[H] = -1$ along the trajectory.

D. Thermodynamic uncertainty relations

Equipped with the above definitions, one can easily verify that the variance in estimating the inverse temperature parameter β can be expressed by the geometrical relation

$$\text{Var}_\psi[B] = \frac{1}{4}g^{ab}\nabla_a\beta\nabla_b\beta \quad (14)$$

on the unit sphere \mathcal{S} , where $\nabla_a\beta = \partial\beta/\partial\psi^a$ is the gradient of the temperature estimate β . The essence of formula (14) can be understood as follows. First, recall that β is the expectation of the estimator B_{ab} in the state $\psi^a(\beta)$. Suppose that the state changes rapidly as β changes. Then, the variance in estimating β is small—indeed, this is given by the squared magnitude of the ‘functional derivative’ of β with respect to the state ψ^a . On the other hand, if the state does not change significantly as β changes, then the measurement outcome of an observable is less conclusive in determining the value of β .

The squared length of the gradient vector $\nabla_a\beta$ can be expressed as a sum of squares of orthogonal components. To this end, we choose a new set of orthogonal basis vectors given by the state ψ^a and its higher derivatives. If we let ψ_n^a denote ψ^a for $n = 0$, and for $n > 0$ the component of the derivative $\partial^n\psi^a/\partial\beta^n$ orthogonal to the state ψ^a and its lower order derivatives, then our orthonormal vectors are given by $\hat{\psi}_n^a = \psi_n^a(g_{bc}\psi_n^b\psi_n^c)^{-1/2}$ for $n = 0, 1, 2, \dots$. With this choice of orthonormal vectors, we find that the variance of the estimator B satisfies

$$\text{Var}_\psi[B] \geq \sum_n \frac{(\tilde{B}_{ab}\psi_n^a\psi_n^b)^2}{g_{cd}\psi_n^c\psi_n^d} , \quad (15)$$

for any range of the index n . This follows because the squared magnitude of the vector $\frac{1}{2}\nabla_a\beta = \tilde{B}_{ab}\psi^b$ is greater than or equal to the sum of the squares of its projections onto the basis vectors given by $\hat{\psi}_n^a$ for the specified range of n .

In particular, for $n = 1$ we have $B_{ab}\psi_1^a\psi^b = \frac{1}{2}$ on account of the relation $B_{ab}\psi^a\psi^b = \beta$, and $g_{ab}\psi_1^a\psi_1^b = \frac{1}{4}\Delta H^2$, which follows from the thermal equation (3). Therefore, if we write $\text{Var}_\psi[B] = \Delta\beta^2$, we find for $n = 1$ that the inequality (15) implies the following thermodynamic uncertainty relation:

$$\Delta\beta^2\Delta H^2 \geq 1 , \quad (16)$$

valid along the trajectory consisting of the thermal equilibrium states. The variance $\Delta^2\beta$ here is to be understood in the sense of estimation theory. That is, although the variable

β does not actually fluctuate, as should be clear from the definition of canonical ensemble, there is nonetheless an inevitable lower bound for the variance of the measurement, given by (16), if we wish to estimate the value of the heat bath temperature β . It is worth pointing out that the exposition we have given here is consistent with the view put forward by Mandelbrot [12], who should perhaps be credited with first having introduced an element of modern statistical reasoning into the long-standing debate on the status of temperature fluctuations [13].

Note that, although we only considered the variance $\langle (B - \langle B \rangle)^2 \rangle$ here, the higher order central moments $\mu_n = \langle (B - \langle B \rangle)^n \rangle$ can also be expressed geometrically. This can be seen as follows. First, recall that for any observable F_{ab} with $E_\psi[F] = f$, we have $\nabla_a f = 2\tilde{F}_{ab}\psi^b$ on the unit sphere \mathcal{S} . Therefore, by letting $F = \tilde{B}^n$, we construct the higher central moments in terms of the cosines of the angles between certain gradient vectors, e.g., $4\mu_3 = g^{ab}\nabla_a\mu_2\nabla_b\beta$, $4\mu_4 = g^{ab}\nabla_a\mu_2\nabla_b\mu_2 - 4\mu_2^2$, and so on. In particular, the even order moments are expressible in terms of combinations of the squared lengths of normal vectors to the surfaces of constant central moments of lower order.

E. Random states

Let us return to the consideration of measurements on thermal states, which we now pursue in greater depth. In doing so we shall introduce the idea of a ‘random state’, a concept that is applicable both in clarifying the measurement problem in statistical physics, as well as in providing a useful tool when we consider ensembles. It also turns out that the idea of a random state is helpful in the analysis of conceptual problems in quantum mechanics. Later on when we consider quantum statistical mechanics, we shall have more to say on this.

Suppose we consider a pure thermal state $\psi^a(\beta)$ for some value β of the inverse temperature. We know that this state is given by

$$\psi^a(\beta) = \sum_k p_k^{1/2} u_k^a, \quad (17)$$

where u_k^a is a normalised energy eigenstate with eigenvalue E_k , and p_k is the associated Gibbs probability. After measurement, it is natural to consider the outcome of the measurement to be a *random state* Ψ^a . Thus we consider Ψ^a to be a random variable (indicated by use of a bold font) such that the probability for taking a given eigenstate is

$$\text{Prob}[\Psi^a = u_k^a] = p_k. \quad (18)$$

This way of thinking about the outcome of the measurement process is to some extent complementary to the density matrix approach, though in what follows we shall make it clear what the relationship is.

In particular, the expectation of an observable X_{ab} in the random state Ψ^a is given by averaging over the random states, that is,

$$E_\Psi[X] = X_{ab} \Psi^a \Psi^b. \quad (19)$$

This relation should be interpreted as the specification of a *conditional* expectation, i.e., the conditional expectation of X_{ab} in the random state Ψ^a . Then the associated *unconditional* expectation $E[X] = E[E_{\Psi}[X]]$ is given by

$$E[X] = X_{ab}E[\Psi^a\Psi^b] . \quad (20)$$

However, since $\text{Prob}[\Psi^a = u_k^a] = p_k$, it should be evident that

$$E[\Psi^a\Psi^b] = \rho^{ab} , \quad (21)$$

where the density matrix ρ^{ab} is defined by

$$\rho^{ab} = \sum_k p_k u_k^a u_k^b . \quad (22)$$

Thus, the unconditional expectation of the random variable X_{ab} is given, as noted earlier, by $E[X] = X_{ab}\rho^{ab}$. It should be observed, however, that here we are not emphasising the role of the density matrix ρ^{ab} as representing a ‘state’, but rather its role in summarising information relating to the random state Ψ^a .

The feature that distinguishes the density matrix in this analysis is that it is fully sufficient for the characterisation of unconditional statistics relating to the observables and states under consideration. This point is clearly illustrated when we calculate the variance of a random variable X_{ab} in a random state Ψ^a . Such a situation arises if we want to discuss the uncertainties arising in the measurement of an observable X_{ab} for an ensemble. In this case the system we have in mind is a large number of identical, independent particles, each of which is in a definite energy eigenstate, where the distribution of the energy is given according to the Gibbs distribution. One might take this as an elementary model for a classical gas. Then the distribution of the ensemble can be described in terms of a random state Ψ^a . Note that here the interpretation is slightly different from what we had considered before (the random outcome of a measurement for an isolated system), though it will be appreciated that the relation of these two distinct interpretations is of considerable interest for physics and statistical theory alike.

The conditional variance of the observable X_{ab} in the random state Ψ^a is given by

$$\text{Var}_{\Psi}[X] = X_{ab}X_c^b\Psi^a\Psi^c - (X_{ab}\Psi^a\Psi^b)^2 . \quad (23)$$

The average over the different values of Ψ^a then gives us

$$E[\text{Var}_{\Psi}[X]] = X_{ab}X_c^b\rho^{ac} - X_{ab}X_{cd}\rho^{abcd} , \quad (24)$$

where ρ^{ab} is, as before, the density matrix (21), and ρ^{abcd} is a certain higher moment of Ψ^a , defined by

$$\rho^{abcd} = E[\Psi^a\Psi^b\Psi^c\Psi^d] . \quad (25)$$

The appearance of this higher order analogue of the density matrix may be surprising, though it is indeed a characteristic feature of conditional probability. However, the unconditional variance of X is not given simply by the expectation $E[\text{Var}_{\Psi}[X]]$, but rather (see, e.g., [14]) by the *conditional variance formula*

$$\text{Var}[X] = E[\text{Var}_{\Psi}[X]] + \text{Var}[E_{\Psi}[X]] . \quad (26)$$

For the second term we have

$$\begin{aligned} \text{Var}[E_{\Psi}[X]] &= E[(X_{ab}\Psi^a\Psi^b)^2] - (E[X_{ab}\Psi^a\Psi^b])^2 \\ &= X_{ab}X_{cd}\rho^{abcd} - (X_{ab}\rho^{ab})^2 , \end{aligned} \quad (27)$$

which also involves the higher moment ρ^{abcd} . The terms in (26) involving ρ^{abcd} then cancel, and we are left with $\text{Var}[X] = X_{ab}X_c^b\rho^{ac} - (X_{ab}\rho^{ab})^2$ for the unconditional variance, which, as indicated earlier, only involves the density matrix. It follows that the random state approach does indeed reproduce the earlier density matrix formulation of our theory, though the role of the density matrix is somewhat diminished. In other words, whenever conditioning is involved, it is the set of totally symmetric tensors

$$\rho^{ab\cdots cd} = E[\Psi^a\Psi^b\cdots\Psi^c\Psi^d] \quad (28)$$

that plays the fundamental role, although it suffices to consider the standard density matrix ρ^{ab} when conditioning is removed.

All this is worth having in mind later when we turn to quantum statistical mechanics, where the considerations we have developed here in a classical context reappear in a new light. We want to de-emphasise the role of the density matrix, not because there is anything wrong *per se* with the use of the density matrix in an appropriate context, but rather for two practical reasons. First of all, when we want to consider conditioning, exclusive attention on the density matrix hampers our thinking, since, as we have indicated, higher moments of the random state vector also have a role to play. Second, when we go to consider generalisations of quantum mechanics, such as the nonlinear theories of the Kibble-Weinberg type [15,16], or stochastic theories of the type considered by Gisin, Percival, and others [17,18], the density matrix is either an ill formulated concept, or plays a diminished role. We shall return to this point for further discussion when we consider quantum statistical mechanics in Section 4.

III. STATISTICAL PHASE SPACE

A. Projective space and probabilities

To proceed further it will be useful to develop a formalism for the algebraic treatment of real projective geometry, with a view to its probabilistic interpretation in the context of classical statistical mechanics. Let Z^a be coordinates for $(n+1)$ -dimensional real Hilbert space \mathcal{H}^{n+1} . Later, when we consider quantum theory from a real point of view we shall double this dimension. In the Hilbert space description of classical probabilities, the normalisation condition is written $g_{ab}Z^aZ^b = 1$. However, this normalisation is physically irrelevant since the expectation of an arbitrary operator F_{ab} is defined by the ratio

$$\langle F \rangle = \frac{F_{ab}Z^aZ^b}{g_{cd}Z^cZ^d} . \quad (29)$$

Therefore, the physical state space is not the Hilbert space \mathcal{H} , but the space of equivalence classes obtained by identifying the points $\{Z^a\}$ and $\{\lambda Z^a\}$ for all $\lambda \in \mathbf{R} - \{0\}$. In this way, we ‘gauge away’ the irrelevant degree of freedom. The resulting space is the real projective n -space RP^n , the space of rays through the origin in \mathcal{H}^{n+1} . Thus, two points X^a and Y^a in \mathcal{H}^{n+1} are equivalent in RP^n if they are proportional, i.e., $X^{[a}Y^{b]} = 0$.

The coordinates Z^a (excluding $Z^a = 0$) can be used as homogeneous coordinates for points of RP^n . Clearly Z^a and λZ^a represent the same point in RP^n . In practice one treats the homogeneous coordinates as though they define points of \mathcal{H}^{n+1} , with the stipulation that the allowable operations of projective geometry are those which transform homogeneously under the rescaling $Z^a \rightarrow \lambda Z^a$.

A *prime*, or $(n - 1)$ -plane in RP^n consists of a set of points Z^a which satisfy a linear equation $P_a Z^a = 0$, where we call P_a the homogeneous coordinates of the prime. Clearly, P_a and λP_a determine the same prime. Therefore, a prime in RP^n is an RP^{n-1} , and the set of all primes in RP^n is itself an RP^n (the ‘dual’ projective space). In particular, the metric g_{ab} on \mathcal{H}^{n+1} can be interpreted in the projective space as giving rise to a nonsingular *polarity*, that is, an invertible map from points to hyperplanes in RP^n of codimension one. This map is given by $P^a \rightarrow P_a := g_{ab} P^b$. See reference [19] for further discussion of some of the geometric operations employed here.

If a point P^a in RP^n corresponds to a probability state, then its negation $\neg P^a$ is the hyperplane $P_a Z^a = 0$. To be more precise, we take P^a as describing the probability state for a set of events. Then, all the probability states corresponding to the complementary events lie in the prime $P_a Z^a = 0$. Thus, the points Z^a on this plane are precisely the states that are orthogonal to the original state P^a . The intuition behind this is as follows. Two states P^a and Q^a are orthogonal if and only if any event which in the state P^a (resp. Q^a) has a positive probability is assigned zero probability by the state Q^a (resp. P^a). This is the sense in which orthogonal states are ‘complementary’. For any state P^a , the plane consisting of all points Z^a such that $P_a Z^a = 0$ is the set of states complementary to P^a in this sense.

Distinct states X^a and Y^a are joined by a real projective line represented by the skew tensor $L^{ab} = X^{[a}Y^{b]}$. The points on this line are the various real superpositions of the original two states. The intersection of the line L^{ab} with the plane R_a is given by $S^a = L^{ab} R_b$. Clearly S^a lies on the plane R_a , since $R_a S^a = 0$ on account of the antisymmetry of L^{ab} .

The hyperplanes that are the negations (polar planes) of two points P^a and Q^a intersect the joining line $L^{ab} = P^{[a}Q^{b]}$ at a pair of points \tilde{P}^a and \tilde{Q}^a respectively. That is, if $P_a = g_{ab} P^b$ is the coordinate of a plane and L^{ab} represents a line in RP^n , then the point of intersection is given by $\tilde{P}^a = L^{ab} P_b$, and similarly $\tilde{Q}^a = L^{ab} Q_b$. The projective cross ratio between these four points $P^a, Q^a, \tilde{P}^a = P^a(Q^b P_b) - Q^a(P^b P_b)$ and $\tilde{Q}^a = P^a(Q^b Q_b) - Q^a(P^b Q_b)$, given by $P^a \tilde{Q}_a Q^b \tilde{P}_b / P^c \tilde{P}_c Q^d \tilde{Q}_d$, reduces, after some algebra, to the following simple expression:

$$\kappa = \frac{(P^a Q_a)^2}{P^b P_b Q^c Q_c}, \quad (30)$$

which can be interpreted as the transition probability between P^a and Q^a . It is interesting to note that this formula has an analogue in quantum mechanics [5].

The projective line L^{ab} can also be viewed as a circle, with \tilde{P}^a and \tilde{Q}^a antipodal to the points P^a and Q^a . In that case, the cross ratio κ is $\frac{1}{2}(1 + \cos \theta) = \cos^2(\theta/2)$ where θ defines the angular distance between P^a and Q^a , in the geometry of RP^n . We note that θ is, in

fact, *twice* the angle made between the corresponding Hilbert space vectors, so orthogonal states are maximally distant from one another.

Now suppose we let the two states P^a and Q^a approach one another. In the limit the resulting formula for their second-order infinitesimal separation determines the natural line element on real projective space. This can be obtained by setting $P^a = Z^a$ and $Q^a = Z^a + dZ^a$ in (30), while replacing θ with the small angle ds in the expression $\frac{1}{2}(1 + \cos \theta)$, retaining terms of the second order in ds . Explicitly, we obtain

$$ds^2 = 4 \left[\frac{dZ^a dZ_a}{Z^b Z_b} - \frac{(Z^a dZ_a)^2}{(Z^b Z_b)^2} \right]. \quad (31)$$

Note that this metric [20] is related to the metric on the sphere \mathcal{S}^n in \mathcal{H}^{n+1} , except that in the case of the sphere one does not identify opposite points.

We now consider the case where the real projective space is the state space of classical statistical mechanics. If we write $\psi^a(\beta)$ for the trajectory of thermal state vectors, as discussed in Section 2, then we can regard $\psi^a(\beta)$, for each value of β , as representing homogeneous coordinates for points in the state space RP^n . Since $\psi^a(\beta)$ satisfies (3) it follows that the line element along the curve $\psi^a(\beta)$ is given by $ds^2 = \langle \tilde{H}^2 \rangle d\beta^2$, which can be identified with the Fisher-Rao metric induced on the thermal trajectory by virtue of its embedding in RP^n . This follows by insertion of the thermal equation (3) into expression (31) for the natural spherical metric on RP^n .

B. The projective thermal equations

Let us write H_{ab} for the symmetric Hamiltonian operator, and $\psi^a(\beta)$ for the one-parameter family of thermal states. Then in our notation the thermal equation is $d\psi^b = -\frac{1}{2}\tilde{H}_c^b \psi^c d\beta$. However, we are concerned with this equation only inasmuch as it supplies information about the evolution of the state of the system, i.e., its motion in RP^n . We are interested therefore primarily in the *projective* thermal equation, given by

$$\psi^{[a} d\psi^{b]} = -\frac{1}{2}\psi^{[a} \tilde{H}_c^{b]} \psi^c d\beta, \quad (32)$$

obtained by skew-symmetrising the thermal equation (3) with ψ^a . Equation (32) defines the equilibrium thermal trajectory of a statistical mechanical system in the proper state space.

The thermal equation generates a Hamiltonian gradient flow on the state manifold. This can be seen as follows. First, recall that a physical observable F is associated with a symmetric operator F_{ab} , and the set of such observables form a vector space of dimension $\frac{1}{2}(n+1)(n+2)$. Such observables are determined by their diagonal matrix elements, which are real valued functions on RP^n of the form $F(\psi^a) = F_{ab}\psi^a\psi^b/\psi^c\psi_c$. In particular, we are interested in the Hamiltonian function $H(\psi^a)$. Then, by a direct substitution, we find that the vector field $H^a = g^{ab}\partial H/\partial\psi^b$ associated with the Hamiltonian function H takes the form $H^a = 2\tilde{H}_b^a\psi^b/\psi^c\psi_c$. Therefore, we can write the differential equation for the thermal state trajectory in \mathcal{H}^{n+1} in the form

$$\frac{d\psi^a}{d\beta} = -\frac{1}{4}g^{ab}\nabla_b H. \quad (33)$$

By projecting this down to RP^n , we obtain

$$\psi^{[a}d\psi^{b]} = -\frac{1}{4}\psi^{[a}\nabla^{b]}Hd\beta, \quad (34)$$

where $\nabla^a H = g^{ab}\nabla_b H$. From this we can then calculate the line element to obtain

$$\begin{aligned} ds^2 &= 8\psi^{[a}d\psi^{b]}\psi_{[a}d\psi_{b]}/(\psi^c\psi_c)^2 \\ &= \frac{1}{2}\psi^{[a}\nabla^{b]}H\psi_{[a}\nabla_{b]}Hd\beta^2 \\ &= \frac{1}{4}\nabla^b H\nabla_b H\psi^a\psi_a d\beta^2 \\ &= \langle \tilde{H}^2 \rangle d\beta^2, \end{aligned} \quad (35)$$

which establishes the result we noted earlier.

The critical points of the Hamiltonian function are the fixed points in the state space associated with the gradient vector field $g^{ab}\nabla_b H$. In the case of Hamilton's equations the fixed points are called stationary states. In the Hilbert space \mathcal{H}^{n+1} these are the points corresponding to the energy eigenstates u_k^a given by $H_b^a u_k^b = E_k u_k^a$ (in the general situation with distinct energy eigenvalues E_k , $k = 0, 1, \dots, n$). Therefore, in the projective space RP^n we have a set of fixed points corresponding to the states u_k^a , and the thermal states are obtained by superposing these points with an appropriate set of coefficients given by the Boltzmann weights. Since these coefficients are nonzero at finite temperature, it should be clear that the thermal trajectories do not intersect any of these fixed points.

In particular, the infinite temperature ($\beta = 0$) thermal state $\psi^a(0)$ is located at the centre point of \mathcal{S}_+ in \mathcal{H}^{n+1} . The distances from $\psi^a(0)$ to the various energy eigenstates are all equal. This implies that in RP^n the cross ratios between the fixed points and the state $\psi^a(0)$ are equal. Therefore, we can single out the point $\psi^a(0)$ as an initial point, and form a geodesic hypersphere in RP^n . All the fixed points of the Hamiltonian vector field H^a lie on this sphere. Since the cross ratios between the fixed points are also equal (i.e., maximal), these fixed points form a regular simplex on the sphere. The thermal trajectory thus commences at $Z^a(0)$, and asymptotically approaches a fixed point associated with the lowest energy eigenvalue E_0 , as $\beta \rightarrow \infty$.

If we take the orthogonal prime of the thermal state for any finite β , i.e., $-\psi^a(\beta)$, then the resulting hyperplane clearly does not contain any of the fixed points. On the other hand, if we take the orthogonal prime of any one of the fixed points u_k^a , then the resulting hyperplane includes a sphere of codimension one where all the other fixed points lie. This sphere is given by the intersection of the original hypersphere surrounding $\psi^a(0)$ with the orthogonal prime of the given excluded fixed point. There is a unique prime containing n general points in RP^n . It is worth noting, therefore, that if we choose n general points given by u_j^a ($j = 0, 1, \dots, n; j \neq k$), then there is a unique solution, up to proportionality, of the n linear equations $X_a u_0^a = 0, \dots, X_a u_n^a = 0$. The solution is then given by $X^a = u_k^a$, where $u_k^a = \epsilon_{bc\dots d}^a u_0^b u_1^c \dots u_n^d$. Here, $\epsilon_{ab\dots c} = \epsilon_{[ab\dots c]}$, with $n+1$ indices, is the totally skew tensor determined up to proportionality.

It would be interesting to explore whether the orientability characteristics of RP^n lead to any physical consequences. There may be a kind of purely classical 'spin-statistics' relation in the sense that the state space of half-integral spins are associated with a topological invariant, while the state space for even spins are not.

C. Hamiltonian flows and projective transformations

We have seen in Section 3.B how the thermal trajectory of a statistical mechanical system is generated by the gradient flow associated with the Hamiltonian function. In other words, for each point on the state manifold we form the expectation of the Hamiltonian in that state. This gives us a global function on the state manifold, which we call the Hamiltonian function. Next, we take the gradient of this function, and raise the index by use of the natural metric to obtain a vector field. This vector field is the generator of the thermal trajectories.

It is interesting to note that there is a relation between the geometry of such vector fields and the global symmetries of the state manifold. In particular, we shall show below that gradient vector fields generated by observables on the state manifold can also be interpreted as the generators of *projective transformations*. A projective transformation on a Riemannian manifold is an automorphism that maps geodesics onto geodesics. In the case of a real projective space endowed with the natural metric, the general such automorphism is generated, as we shall demonstrate, by a vector field that is expressible in the form of a sum of a Killing vector and a gradient flow associated with an observable function.

To pursue this point further we develop some differential geometric aspects of the state manifold. We consider RP^n now to be a differential manifold endowed with the natural spherical metric g_{ab} . Here, bold upright indices signify local tensorial operations in the tangent space of this manifold. Thus we write $\nabla_{\mathbf{a}}$ for the covariant derivative associated with g_{ab} , and for an arbitrary vector field $V^{\mathbf{a}}$ we define the Riemann tensor $R_{\mathbf{abc}}{}^{\mathbf{d}}$ according to the convention

$$\nabla_{[\mathbf{a}}\nabla_{\mathbf{b}]}V^{\mathbf{c}} = \frac{1}{2}R_{\mathbf{abd}}{}^{\mathbf{c}}V^{\mathbf{d}} . \quad (36)$$

It follows that $R_{\mathbf{abcd}} := R_{\mathbf{abc}}{}^{\mathbf{e}}g_{\mathbf{de}}$ satisfies $R_{\mathbf{abcd}} = R_{[\mathbf{ab}][\mathbf{cd}]}$, $R_{\mathbf{abcd}} = R_{\mathbf{cdab}}$, and $R_{[\mathbf{abc}]\mathbf{d}} = 0$. In the case of RP^n with the natural metric

$$ds^2 = \frac{8Z^{[a}dZ^{b]}Z_{[a}dZ_{b]}}{\lambda(Z^cZ_c)^2} , \quad (37)$$

where Z^a are homogeneous coordinates and λ is a scale factor (set to unity in the preceding analysis), the Riemann tensor is given by

$$R_{\mathbf{abcd}} = \lambda(g_{\mathbf{ac}}g_{\mathbf{bd}} - g_{\mathbf{bc}}g_{\mathbf{ad}}) . \quad (38)$$

Now we turn to consider projective transformations on RP^n . First we make a few general remarks about projective transformations on Riemannian manifolds [21]. Suppose we have a Riemannian manifold with metric g_{ab} and we consider the effect of dragging the metric along the integral curves of a vector field $\xi^{\mathbf{a}}$. For an infinitesimal transformation we have

$$g_{\mathbf{ab}} \rightarrow \hat{g}_{\mathbf{ab}} = g_{\mathbf{ab}} + \epsilon \mathcal{L}_{\xi}g_{\mathbf{ab}} , \quad (39)$$

where $\mathcal{L}_{\xi}g_{\mathbf{ab}} = 2\nabla_{(\mathbf{a}}\xi_{\mathbf{b})}$ is the Lie derivative of the metric ($\epsilon \ll 1$). The Levi-Civita connection $\hat{\nabla}_{\mathbf{a}}$ associated with $\hat{g}_{\mathbf{ab}}$ is then defined by a tensor $Q_{\mathbf{ab}}{}^{\mathbf{c}}$, symmetric on its lower indices, such that

$$\hat{\nabla}_{\mathbf{a}} V^{\mathbf{c}} = \nabla_{\mathbf{a}} V^{\mathbf{c}} + \epsilon Q_{\mathbf{ab}}^{\mathbf{c}} V^{\mathbf{b}} , \quad (40)$$

for an arbitrary vector field $V^{\mathbf{a}}$, and $\hat{\nabla}_{\mathbf{a}} \hat{g}_{\mathbf{bc}} = 0$. A familiar line of argument based on the fundamental theorem of Riemannian geometry then shows that

$$Q_{\mathbf{ab}}^{\mathbf{c}} = \frac{1}{2} \hat{g}^{\mathbf{cd}} (\nabla_{\mathbf{a}} \hat{g}_{\mathbf{bd}} + \nabla_{\mathbf{b}} \hat{g}_{\mathbf{ad}} + \nabla_{\mathbf{d}} \hat{g}_{\mathbf{ab}}) , \quad (41)$$

where $\hat{g}^{\mathbf{ab}}$ is the inverse of $\hat{g}_{\mathbf{ab}}$, given to first order in ϵ here by $\hat{g}^{\mathbf{ab}} = g^{\mathbf{ab}} - 2\epsilon \nabla^{(\mathbf{a}} \xi^{\mathbf{b})}$. A short calculation making use of (36) then shows that to first order we have:

$$Q_{\mathbf{ab}}^{\mathbf{c}} = \nabla_{(\mathbf{a}} \nabla_{\mathbf{b})} \xi^{\mathbf{c}} + R_{(\mathbf{ab})\mathbf{d}}^{\mathbf{c}} \xi^{\mathbf{d}} . \quad (42)$$

Now, suppose the vector field $U^{\mathbf{a}}$ satisfies the geodesic equation, which we write in the form $(U^{\mathbf{a}} \nabla_{\mathbf{a}} U^{\mathbf{b}}) U^{\mathbf{c}} = 0$. It should be apparent that $U^{\mathbf{a}}$ is geodesic with respect to the transformed metric $\hat{g}_{\mathbf{ab}}$ iff $(U^{\mathbf{a}} \hat{\nabla}_{\mathbf{a}} U^{\mathbf{b}}) U^{\mathbf{c}} = 0$, or equivalently, $U^{\mathbf{a}} U^{\mathbf{b}} Q_{\mathbf{ab}}^{\mathbf{c}} U^{\mathbf{d}} = 0$. However, this relation is satisfied for *all* geodesic vector fields iff there exists a vector field $\phi_{\mathbf{a}}$ such that $Q_{\mathbf{ab}}^{\mathbf{c}} = \delta_{(\mathbf{a}}^{\mathbf{c}} \phi_{\mathbf{b})}$. Hence we conclude that the vector field $\xi^{\mathbf{a}}$ is the generator of a projective transformation, mapping geodesics to geodesics, iff there exists a vector field $\phi_{\mathbf{a}}$ such that

$$\nabla_{(\mathbf{a}} \nabla_{\mathbf{b})} \xi^{\mathbf{c}} + R_{(\mathbf{ab})\mathbf{d}}^{\mathbf{c}} \xi^{\mathbf{d}} = \delta_{(\mathbf{a}}^{\mathbf{c}} \phi_{\mathbf{b})} \quad (43)$$

On the other hand, it follows as a direct consequence of (36) and the identity $R_{[\mathbf{abc}]\mathbf{d}} = 0$ that for *any* vector field $\xi^{\mathbf{a}}$ we have

$$\nabla_{[\mathbf{a}} \nabla_{\mathbf{b}]} \xi^{\mathbf{c}} - R_{[\mathbf{ab}]\mathbf{d}}^{\mathbf{c}} \xi^{\mathbf{d}} = 0 . \quad (44)$$

Hence, combining (43) and (44) we obtain

$$\nabla_{\mathbf{a}} \nabla_{\mathbf{b}} \xi^{\mathbf{c}} + R_{\mathbf{bad}}^{\mathbf{c}} \xi^{\mathbf{d}} = \delta_{(\mathbf{a}}^{\mathbf{c}} \phi_{\mathbf{b})} , \quad (45)$$

as a condition equivalent to (43) for the existence of a projective transformation on the given manifold. This form of the condition is particularly useful for calculations (see, e.g., [21]).

In particular, if $\xi^{\mathbf{a}}$ is a Killing vector, so $\nabla_{(\mathbf{a}} \xi_{\mathbf{b})} = 0$, then (45) is automatically satisfied, with $\phi_{\mathbf{a}} = 0$ since as a consequence of (36) a Killing vector necessarily satisfies $\nabla_{\mathbf{a}} \nabla_{\mathbf{b}} \xi^{\mathbf{c}} + R_{\mathbf{bad}}^{\mathbf{c}} \xi^{\mathbf{d}} = 0$. Thus, a Killing symmetry generates a projective transformation. For example, in the case of RP^n we have the symmetries of the projective orthogonal group.

Now we derive an integrability condition for (45) that will be helpful in our analysis of the remaining projective transformations on RP^n , apart from those associated with Killing symmetries. Returning to the transformation (39), we consider the curvature $\hat{R}_{\mathbf{abc}}^{\mathbf{d}}$ associated with the metric $\hat{g}_{\mathbf{ab}}$. Clearly, to first order in ϵ this is given by

$$\hat{R}_{\mathbf{abc}}^{\mathbf{d}} = R_{\mathbf{abc}}^{\mathbf{d}} + \epsilon \mathcal{L}_{\xi} R_{\mathbf{abc}}^{\mathbf{d}} . \quad (46)$$

On the other hand, it is well known (see, e.g., [22]) that for a change of connection given by (40) the corresponding change in the curvature is

$$\frac{1}{2} \hat{R}_{\mathbf{abc}}^{\mathbf{d}} = \frac{1}{2} R_{\mathbf{abc}}^{\mathbf{d}} + \epsilon \nabla_{[\mathbf{a}} Q_{\mathbf{b}]\mathbf{c}}^{\mathbf{d}} + \epsilon^2 Q_{\mathbf{r}[\mathbf{a}}^{\mathbf{d}} Q_{\mathbf{b}]\mathbf{c}}^{\mathbf{r}} . \quad (47)$$

Thus by consideration of terms to first order in ϵ we deduce that

$$\frac{1}{2}\mathcal{L}_\xi R_{\mathbf{abc}}{}^{\mathbf{d}} = \nabla_{[\mathbf{a}}Q_{\mathbf{b}]c}{}^{\mathbf{d}} , \quad (48)$$

with $Q_{\mathbf{ab}}^{\mathbf{c}}$ given as in equation (42). This relation holds for any infinitesimal transformation of the form (39) generated by a given vector field $\xi^{\mathbf{a}}$. In the case of a projective transformation, for which $Q_{\mathbf{ab}}^{\mathbf{c}} = \delta_{\mathbf{a}}^{\mathbf{c}}\phi_{\mathbf{b}}$, it follows, therefore, after some elementary rearrangement, that

$$\mathcal{L}_\xi R_{\mathbf{abc}}{}^{\mathbf{d}} + \delta_{[\mathbf{a}}^{\mathbf{d}}\nabla_{\mathbf{b}]} \phi_{\mathbf{c}} - \nabla_{[\mathbf{a}}\phi_{\mathbf{b}]} \delta_{\mathbf{c}}^{\mathbf{d}} = 0 . \quad (49)$$

This is the desired integrability condition. In the case of a space of constant curvature we have (38) and hence

$$\mathcal{L}_\xi R_{\mathbf{abc}}{}^{\mathbf{d}} = 2\lambda\nabla_{(\mathbf{a}}\xi_{\mathbf{c})}\delta_{\mathbf{b}}^{\mathbf{d}} - 2\lambda\nabla_{(\mathbf{b}}\xi_{\mathbf{c})}\delta_{\mathbf{a}}^{\mathbf{d}} . \quad (50)$$

Inserting this relation into equation (49) we obtain $A_{\mathbf{ac}}g_{\mathbf{bd}} - A_{\mathbf{bc}}g_{\mathbf{ad}} = \nabla_{[\mathbf{a}}\phi_{\mathbf{b}]}g_{\mathbf{cd}}$, where $A_{\mathbf{ab}} = 2\lambda\nabla_{(\mathbf{a}}\xi_{\mathbf{b})} - \nabla_{\mathbf{a}}\phi_{\mathbf{b}}$. This implies $A_{\mathbf{ab}} = 0$ and $\nabla_{[\mathbf{a}}\xi_{\mathbf{b}]} = 0$. For otherwise, we could contract $A_{\mathbf{ab}}$ with a vector $B^{\mathbf{c}}$ to obtain $C_{\mathbf{a}}g_{\mathbf{bd}} - C_{\mathbf{b}}g_{\mathbf{ad}} = \nabla_{[\mathbf{a}}\phi_{\mathbf{b}]}B_{\mathbf{d}}$ where $C_{\mathbf{a}} = A_{\mathbf{ac}}B^{\mathbf{c}}$. But this would imply that $g_{\mathbf{ab}}$ is at most of rank two. It follows therefore that $\phi_{\mathbf{a}}$ is the gradient of a scalar, and that $\xi^{\mathbf{a}}$ is necessarily of the form

$$\xi^{\mathbf{a}} = \eta_{\mathbf{a}} + g^{\mathbf{ab}}\nabla_{\mathbf{b}}H , \quad (51)$$

for some function $H(x)$ on the state manifold, where $\eta_{\mathbf{a}}$ is a Killing vector, and $\phi_{\mathbf{a}} = 2\lambda\nabla_{\mathbf{a}}H$. Returning this information to equation (43), we conclude that

$$\nabla_{(\mathbf{a}}\nabla_{\mathbf{b}}\nabla_{\mathbf{c})}H = 2\lambda g_{(\mathbf{ab}}\nabla_{\mathbf{c})}H . \quad (52)$$

In other words, $K_{\mathbf{ab}} := \nabla_{\mathbf{a}}\nabla_{\mathbf{b}}H - 2\lambda g_{\mathbf{ab}}H$ is a Killing tensor: $\nabla_{(\mathbf{a}}K_{\mathbf{bc})} = 0$. However, this is the defining equation on RP^n for observable functions, i.e., functions of the form

$$H(x) = \frac{H_{ab}Z^aZ^b}{g_{cd}Z^cZ^d} , \quad (53)$$

in terms of homogeneous coordinates on RP^n . Thus, we conclude that the generator of a projective transformation on the state space is given by the sum of a Killing vector and the gradient of an observable function.

We remark, incidentally, that if the dimension n of the manifold is finite, then we can form the trace of (45), which gives $\nabla_{\mathbf{a}}(\nabla^2H) = \lambda(n+1)\nabla_{\mathbf{a}}H$, from which we conclude that up to an additive constant, $H(x)$ is an eigenfunction of the Laplacian operator, i.e., $\nabla^2H = \lambda(n+1)H$. The only global solutions of this equation on RP^n are necessarily of the form (53). Such functions also satisfy (52). In the case of infinite dimension the trace is not necessarily a valid operation, and we lose the representation of observable functions as eigenfunctions of the Laplacian. On the other hand, our characterisation of projective transformations as vector fields of the form (51), given by the sum of a Killing vector and the gradient of a scalar $H(x)$ satisfying (52) and (53), is entirely general, valid in both finite and infinite dimensions. Setting $\lambda = 1$, we see therefore that thermal states are

characterised by vector fields on RP^n of the form $\xi^a = g^{ab}\nabla_b H$, where $H(x)$ satisfies $\nabla_{(a}\nabla_b\nabla_c)H = 2g_{(ab}\nabla_{c)}H$.

It is worthwhile noting that among the trajectories generated by the Hamiltonian gradient flow, the thermal trajectory is singled out on account of the special initial state corresponding to $\beta = 0$. This state is, of course, completely determined by the specification of the Hamiltonian function. In particular, for a given Hamiltonian H_{ab} there are in general $n + 1$ fixed points in the state manifold. The initial state for the thermal trajectory is then uniquely determined by the condition that it is equidistant from these $n + 1$ fixed points.

IV. QUANTUM STATISTICAL MECHANICS

A. The quantum phase space

In the algebraic approach to quantum theory one generally works with the totality of mixed states, while regarding pure states as extremal elements of this convex set [10], and in the Heisenberg picture for dynamics. The approach we adopt here, however, will focus primarily on the system of pure states, and for dynamics we shall work in the Schrödinger picture. This is consistent with the point of view we put forward earlier in Section 2 in connection with the discussion of classical thermal states.

Given a complex Hilbert space \mathcal{H}_C^{n+1} with complex coordinates Z^α representing pure quantum states, we identify the state vector Z^α with its complex multiples λZ^α , $\lambda \in \mathbf{C} - \{0\}$, to obtain the complex projective space CP^n . Here we use Greek indices ($\alpha = 0, 1, \dots, n$) for vectors in \mathcal{H}_C^{n+1} . Following the line of argument indicated earlier in the case of the thermal state space, we regard CP^n as the true ‘state space’ of quantum mechanics. The status of density matrices will be discussed shortly. Our goal now is to build up the necessary projective geometric machinery appropriate for the consideration of quantum mechanics.

Suppose we regard the state vector Z^α as representing homogeneous coordinates on the projective Hilbert space. The complex conjugation of the state vector Z^α is written \bar{Z}_α , with the corresponding Hermitian inner product $Z^\alpha \bar{Z}_\alpha$. The complex conjugate of a point P^α in CP^n is the hyperplane (prime) $\bar{P}_\alpha Z^\alpha = 0$. The points on this plane are the states that are orthogonal to the original state P^α , and we denote this $\neg P^\alpha$ in the sense of the probability rules discussed above. Thus, CP^n is equipped with a Hermitian correlation, i.e., a complex conjugation operation that maps points to hyperplanes of codimension one, that is, CP^{n-1} .

Distinct points X^α and Y^α are joined by a complex projective line $L^{\alpha\beta} = X^{[\alpha}Y^{\beta]}$, representing the various complex, quantum mechanical superpositions of the original two states. The quantum mechanical transition probability between two states X^α and Y^α is then given by the cross ratio

$$\kappa = \frac{X^\alpha \bar{Y}_\alpha Y^\beta \bar{X}_\beta}{X^\gamma \bar{X}_\gamma Y^\delta \bar{Y}_\delta}. \quad (54)$$

More precisely, we recall that if the system is in the state Y^α and a measurement is made to see if the system is in the state X^α (corresponding to a measurement of the observable represented by the projection operator $X^\alpha \bar{X}_\beta / X^\gamma \bar{X}_\gamma$), then κ is the probability that the result of the measurement is affirmative. If we set $\kappa = \cos^2(\theta/2)$ and set $X^\alpha = Z^\alpha$ and

$Y^\alpha = Z^\alpha + dZ^\alpha$, retaining terms to second order, we recover the natural unitary-invariant metric on complex projective space, known as the Fubini-Study metric [20], given by

$$ds^2 = 8(Z^\alpha \bar{Z}_\alpha)^{-2} Z^{[\alpha} dZ^{\beta]} \bar{Z}_{[\alpha} d\bar{Z}_{\beta]} . \quad (55)$$

Now suppose we write H_β^α for the Hamiltonian operator, assumed Hermitian. Then for the Schrödinger equation we have

$$\frac{dZ^\alpha}{dt} = iH_\beta^\alpha Z^\beta . \quad (56)$$

However, we are more interested in the *projective* Schrödinger equation, given by

$$Z^{[\alpha} dZ^{\beta]} = iZ^{[\alpha} H_\gamma^{\beta]} Z^\gamma dt , \quad (57)$$

which eliminates the superfluous degree of freedom associated with the direction of Z^α . This equation is well defined on the projective state space. Insertion of (57) into (55) gives

$$ds^2 = 4\langle \tilde{H}^2 \rangle dt^2 , \quad (58)$$

which verifies that the velocity of a state along its trajectory in state space is $ds/dt = 2\Delta H$, where $\Delta H = \langle \tilde{H}^2 \rangle^{1/2}$ is the energy uncertainty [23].

An alternative way of looking at this structure is to regard the state manifold CP^n as a real manifold \mathcal{M}^{2n} of dimension $2n$, equipped with a Riemannian metric $g_{\mathbf{ab}}$ and a compatible complex structure $\Omega_{\mathbf{ab}}$. Here again we use bold indices for tensorial operations in the tangent space of \mathcal{M}^{2n} . In this formulation the Schrödinger equation takes the form of a flow $\xi^{\mathbf{a}}$ on \mathcal{M}^{2n} given by

$$\xi^{\mathbf{a}} = \Omega^{\mathbf{ab}} \nabla_{\mathbf{b}} H . \quad (59)$$

Here H is the real function on \mathcal{M}^{2n} given at each point by the expectation of the Hamiltonian operator at that state. It is interesting to note that in the quantum mechanical case the dynamical trajectory is given by a Hamiltonian *symplectic* flow in contrast to the case of statistical mechanics where the relevant flow is given by a Hamiltonian *gradient* flow for the thermal trajectories. Furthermore, we note that in the case of quantum mechanics the flow is necessarily a Killing field, satisfying $\nabla_{(\mathbf{a}} \xi_{\mathbf{b})} = 0$, where $\xi_{\mathbf{a}} = g_{\mathbf{ab}} \xi^{\mathbf{b}}$. In other words, the isometries of the Fubini-Study metric on CP^n can be lifted to $\mathcal{H}_{\mathbf{C}}^{n+1}$ to yield unitary transformations. In the statistical mechanical case, on the other hand, the general projective transformation of the manifold includes both Killing vectors and Hamiltonian gradient flows, but we exclude the former. In both quantum mechanics and statistical mechanics the fixed points of the Hamiltonian operator play a pivotal role.

B. The quantum thermal states

In Section 4.A above we have indicated two useful ways of looking at the quantum state space. The approach via complex homogeneous coordinates is appropriate when the *global* algebraic geometry of the state manifold is of interest. The differential geometric approach is useful when *local* properties of the state manifold are considered [18]. There is a third,

equally important approach, however, which we shall call the ‘real’ approach [6,8,24] to quantum mechanics. In this case we regard the Hilbert space of quantum mechanics as a real vector space of dimension $2n + 2$. The importance of the real approach is that it links up directly with modern notions from probability theory and statistics. It should not therefore be surprising that this approach is also useful in considering the quantum dynamics of thermal states, and indeed in considering the relation between classical and quantum statistical mechanics.

Let us write ξ^a ($a = 1, 2, \dots, 2n + 2$) for the real Hilbert space coordinates corresponding to a quantum mechanical state vector. The other ingredients we have at our disposal are the metric g_{ab} on \mathcal{H}^{2n+2} and a compatible complex structure tensor J_a^b . Therefore, we proceed in two stages. First we regard the state vector ξ^a as homogeneous coordinates for a real projective space RP^{2n+1} of one dimension lower. This, of course, corresponds to factoring out by the freedom $\xi^a \rightarrow \lambda \xi^a$ with λ real and nonvanishing. Then there is another more subtle freedom, corresponding to factoring out by the ‘phase’ freedom. This is given by the transformation

$$\xi^a \rightarrow \frac{1}{2}e^{i\phi}(\xi^a + iJ_b^a\xi^b) + \frac{1}{2}e^{-i\phi}(\xi^a - iJ_b^a\xi^b), \quad (60)$$

where J_b^a is the complex structure, and ϕ is a phase factor. In this way we obtain a map $RP^{2n+1} \rightarrow CP^n$, which is related to the Hopf map $S^{2n+1} \rightarrow CP^n$, and indeed we have the diagram

$$\begin{array}{ccc} \mathcal{H}^{2n+2} & \longrightarrow & S^{2n+1} \\ & \searrow & \downarrow \quad \searrow \\ & & RP^{2n+1} \longrightarrow CP^n \end{array} \quad (61)$$

showing the relation between these various spaces. The point of this line of reasoning is that if we take the classical statistical argument based on the state space RP^{2n+1} , and introduce a complex structure on the real Hilbert space \mathcal{H}^{2n+2} , then the analogous relations for quantum statistical mechanics, and moreover, quantum dynamics of the equilibrium thermal states can be developed by studying the relation between RP^{2n+1} and CP^n .

Now we consider more explicitly the characterisation of thermal states in the context of quantum ‘thermodynamics’. There are a number of features of the Hilbert space based description of conventional statistical mechanics that carry through to the quantum regime, though there are some new features as well. In particular, if we start with the real Hilbert space \mathcal{H}^{2n+2} and the associated state space RP^{2n+1} obtained when we neglect the complex structure on \mathcal{H}^{2n+2} given by quantum theory, then we can apply the theory developed earlier to define thermal trajectories in the space RP^{2n+1} . The main difference arising between the classical and quantum cases is that in the quantum theory the Hamiltonian operator H_{ab} satisfies the Hermitian property

$$H_{ab}J_c^aJ_d^b = H_{cd}, \quad (62)$$

which ensures that it has at most $n + 1$ distinct eigenvalues, unlike the classical situation where there could be $2n + 2$. The thermal equation giving the development of $\psi^a(\beta)$ as the inverse temperature β is changed, is again given by

$$\frac{d\psi^a(\beta)}{d\beta} = \frac{1}{2}\tilde{H}_b^a\psi^b(\beta) , \quad (63)$$

where $\tilde{H}_{ab} = H_{ab} - g_{ab}H_{cd}\psi^c\psi^d/g_{ef}\psi^e\psi^f$.

Another feature that distinguishes the quantum theory is that in general there is a multiplicity of infinite temperature ($\beta = 0$) states, corresponding to different values of the ‘phases’ associated with the state. Topologically, this implies that there is a torus $T^n = (S^1)^n$ in CP^n , the points of which correspond to the $\beta = 0$ state. Thus, for a specified infinite temperature state, i.e., for a given point on T^n , we shall call the solution of (63) a ‘primitive’ thermal state parameterised by a set of phases. In other words, there is a multiplicity of primitive thermal trajectories, each corresponding to a choice of infinite temperature state. In the case of a two state system (e.g., a spin one-half particle) the two energy eigenstates correspond to the north and south poles of the state space $CP^1 \sim S^2$, and the infinite temperature states are given by points on the equator taking different values of the phases. The thermal trajectories are then given by the geodesic segments that join the equator to one of the poles. This example is studied in more detail in Section 4.D.

For many applications we have to consider ensembles of particles for which the starting point of the thermal trajectory is effectively random. In other words, although each particle in the ensemble has a definite thermal trajectory, when we form expectations based on the behaviour of the ensemble we have to average over the manifold of infinite temperature states, i.e., average over the random phase factors. This leads to the well known density matrix based description of thermal states, as we discuss below. Thus once again the density matrix formulation emerges as an essentially secondary construction, although of course it remains essential for many practical applications.

Another characteristic of quantum thermal dynamics is that the thermal trajectories have a time evolution, given by the Schrödinger equation. Written in terms of real state vectors this evolution is given by

$$\psi_\phi^a(t, \beta) = \exp(-tJ_b^a\tilde{H}_c^b)\psi_\phi^c(0, \beta) , \quad (64)$$

for each thermal trajectory, where the variables collectively denoted ϕ label the $\beta = 0$ states. Thus the primitive thermal states are pure states, subject to the usual evolutionary laws.

In our earlier discussion of classical thermal systems, we introduced the idea of a random state. This notion is also useful in a quantum mechanical context. For example, a random state Ψ^a can be used to represent the outcome of a measurement process, or to describe the statistics of an ensemble. In particular, as we indicated earlier, in the case of a quantum thermal state we want to take an ensemble average over all possible infinite temperature states to describe the observed features of ensembles. This point will now be explored further.

Suppose we denote by Φ the system of random phases labelling the $\beta = 0$ states for a given quantum ensemble. The thermal state ψ_Φ^a is thus itself to be viewed as a random state, since Φ is random. Then if X_{ab} is a typical observable, satisfying the Hermitian property, we can write $E_\Phi[X] = X_{ab}\psi_\Phi^a\psi_\Phi^b$ for the conditional expectation of X_{ab} , given Φ . This is in keeping with the line of argument introduced earlier, for classical thermal systems. The ensemble average $E[X]$ is then given by the unconditional expectation

$$E[E_\Phi[X]] = X_{ab}\rho^{ab} , \quad (65)$$

where

$$\rho^{ab}(\beta) = E[\psi_{\Phi}^a \psi_{\Phi}^b] \quad (66)$$

is the Hermitian density matrix characterising the state of the ensemble. Here the unconditional expectation $E[-]$ averages over the phase variables appearing in random state $\psi_{\Phi}^a(0, \beta)$. In this connection it is worth noting that the infinite temperature states have the property that the ensemble average in formula (66) is proportional to the metric for $\beta = 0$, that is, $\rho^{ab}(0) = \frac{1}{2n+2}g^{ab}$, where $2n+2$ is the dimension of the real Hilbert space.

It is interesting to take note of the topological characteristics of the infinite temperature state manifold. As we already remarked, this manifold can be viewed as an n -torus T^n sitting in the state manifold CP^n . This can be seen as follows. First, recall the fact that the $\beta = 0$ states are equidistant from the fixed points u_k^α of the Hamiltonian flow in CP^n , and that these fixed points form a regular simplex. Thus, we would like to find the locus of points Z^α such that the distances between Z^α and u_k^α are equal for all k . Since the points u_k^α are vertices of a regular simplex, we can choose coordinates, without loss of generality, such that $u_0^\alpha = (1, 0, 0, \dots)$, $u_1^\alpha = (0, 1, 0, \dots)$, and so on. Thus, by writing the equidistance condition in terms of the cross ratio explicitly, we find that the locus of $Z^\alpha = (Z^0, Z^1, \dots, Z^n)$ is given by the simultaneous solution of the following system of equations: $Z^0 \bar{Z}_0 = Z^1 \bar{Z}_1, Z^1 \bar{Z}_1 = Z^2 \bar{Z}_2, \dots, Z^n \bar{Z}_n = Z^0 \bar{Z}_0$. We thus deduce that the solution takes the form

$$Z^\alpha = (e^{i\phi_0}, e^{i\phi_1}, \dots, e^{i\phi_n}), \quad (67)$$

for some phase variables ϕ_j . This implies that the topology of the infinite temperature states manifold indeed has a structure of an n -torus $T^n = (S^1)^n$ (one of the $n+1$ phase variables ϕ_j is redundant, and can be scaled away), as claimed earlier. In summary, the fundamental torus $T^n \subset CP^n$ associated with a given Hamiltonian function consists of all those states that are equally distant to the eigenstates of the specified Hamiltonian with respect to the Fubini-Study metric on CP^n . The significance of T^n is that it embodies a geometrical representation of the concept of the extraneous phases over which one has to integrate for many purposes in thermal physics.

C. The equilibrium condition

The property that the density matrix $\rho^{ab}(0)$ at infinite temperature is proportional to the metric is, in fact, closely related to the KMS-condition [10,25], which provides an alternative characterisation of thermal equilibrium states. It is therefore interesting to see how the KMS construction fits into the present description of thermal phenomena, and here we shall briefly develop some of the relevant ideas. First, recall that a general state, or density matrix, can be regarded as a semidefinite map $\rho(A)$ from Hermitian operators A to the real numbers given by $\rho(A) = \text{Tr}(\rho A)$, satisfying $\rho(X\bar{X}) \geq 0$ for all X , and $\rho(g) = 1$, where g is the identity operator. Now the trace operation has the property $\text{Tr}(AB) = \text{Tr}(BA)$. For a general state ρ , on the other hand, clearly we do not have $\rho(AB) = \rho(BA)$. However, the thermal equilibrium states $\rho_\beta(-)$ are characterised by a slightly weakened form of commutation relation, which is the KMS condition.

Suppose for a Hermitian operator A we write $\tau_t A$ for the unitary action of the one-parameter group of time translations given by $\tau_t A = e^{itH} A e^{-itH}$. Then, we can extend this definition to complex time by writing $\tau_{t+i\beta} A = e^{i(t+i\beta)H} A e^{-i(t+i\beta)H}$. The KMS condition on a state ρ is that for all Hermitian operators A and B it should satisfy

$$\rho(\tau_t AB) = \rho(B \tau_{t+i\beta} A) . \quad (68)$$

It is not difficult to see that in finite dimensions this implies that ρ is the thermal state ρ_β discussed above, and that commutivity holds for infinite temperature.

Now let us look at these relations from a real point of view. First we need to consider the action of the one-parameter group of time translations. Suppose we have a real Hilbert space of dimension $2n+2$, with typical elements $\xi^a, \eta^a \in \mathcal{H}$. On \mathcal{H} we also have the metric g_{ab} and the complex structure J_a^b , satisfying $J_a^c J_c^b = -\delta_a^b$ and $J_a^c J_b^d g_{cd} = g_{ab}$. The tensor $\Omega_{ab} = g_{ac} J_b^c$ is thus antisymmetric, and defines a symplectic structure on \mathcal{H} . With these ingredients at hand, we can define the action of the orthogonal and unitary groups on \mathcal{H} .

For a typical element $\xi^a \in \mathcal{H}$ the orthogonal group $O(2n+2)$ consists of transformations $\xi^a \rightarrow O_a^b \xi^b$, satisfying $O_a^c O_b^d g_{cd} = g_{ab}$. The unitary group $U(n+1)$ can then be regarded as a subgroup of $O(2n+2)$, given by transformations $\xi^a \rightarrow U_a^b \xi^b$ satisfying

$$U_a^c U_b^d g_{cd} = g_{ab} , \quad U_a^c U_b^d \Omega_{cd} = \Omega_{ab} . \quad (69)$$

In other words, transformations that preserve both the metric and the symplectic structure on \mathcal{H} . For typical real elements $\xi^a, \eta^a \in \mathcal{H}$ their ordinary real Hilbert space inner product is, of course, given by $\eta^a g_{ab} \xi^b$, which is invariant if ξ^a and η^a are subject to orthogonal transformations. On the other hand, the standard Dirac product $\langle \eta | \xi \rangle$ between two real vectors ξ^a and η^a , given by

$$\langle \eta | \xi \rangle = \frac{1}{2} \eta^a (g_{ab} - i \Omega_{ab}) \xi^b , \quad (70)$$

is invariant only under unitary transformations. This should be clear from the need to preserve the real and imaginary part of (70) separately, which implies (69).

If O_a^b is an orthogonal transformation, then the corresponding inverse element \bar{O}_a^b is given by the transpose $\bar{O}_a^b = g^{bc} g_{ad} O_c^d$, which satisfies $O_b^a \bar{O}_c^b = \delta_c^a$. In the case of a unitary transformation we can, without ambiguity, use the notation $\bar{U}_a^b = g^{bc} g_{ad} U_c^d$ for the conjugate transformation, from which it follows that

$$U_b^a \bar{U}_c^b = \delta_c^a , \quad U_b^a J_c^b \bar{U}_d^c = \delta_d^a \quad (71)$$

are equivalent to (69) in characterising unitary transformations. In particular, (71) can be viewed as saying that the unitary transformations on \mathcal{H} are orthogonal transformations that also preserve the complex structure.

In equation (71) we see the action of the unitary group on the complex structure tensor. More generally, for a typical multi-index tensor A_{bc}^a we define the action of the unitary group by $A_{bc}^a \rightarrow U_a^a A_{b'c'}^{a'} \bar{U}_b^{b'} \bar{U}_c^{c'}$. Here the primed indices merely serve to increase the size of the standard alphabet.

Now suppose we consider one-parameter subgroups of orthogonal group, continuous with the identity. Such transformations are of the form $O_a^b = \exp[t M_{ac} g^{bc}]$ where the tensor

M_{ab} is antisymmetric. For a one-parameter family of unitary transformations we have to specialise further, and require M_{ab} to be of the form $M_{ab} = \Omega_{ac} H_b^c$, where H_{ab} is symmetric and Hermitian. Indeed, if H_{ab} is symmetric, then a necessary and sufficient condition that it should be Hermitian is that $\Omega_{ac} H_b^c$ is antisymmetric. It follows that the general one-parameter group of unitary transformations, continuous with the identity, can be written

$$\xi^a \rightarrow \xi_t^a = \exp [t J_b^a H_c^b] \xi^c . \quad (72)$$

If H_{ab} is the Hamiltonian, then ξ_t^a is the Schrödinger evolution generated from the given initial state. The point of view here continues to be purely ‘real’ in the sense that the complex structure tensor J_b^a is playing the role of the factor of ‘ i ’ in the conventional expression of unitary evolution given in the Dirac notation by $|\xi\rangle \rightarrow e^{itH}|\xi\rangle$. Our goal is to formulate a similar geometrisation for the KMS condition.

For this purpose we need to consider the changes implied for the picture noted above when we go to the Heisenberg representation. Suppose A_{ab} is an observable (symmetric, Hermitian operator) and we consider the evolution of its expectation $A_{ab}\xi^a\xi^b$ under the action of the unitary transformation $\xi^a \rightarrow U_b^a \xi^b$ with $U_b^a = \exp[t J_c^a H_b^c]$. The Heisenberg picture is obtained if we let ξ^a be stationary, and let A_{ab} evolve according to the scheme

$$A_{ab} \rightarrow U_a^c U_b^d A_{cd} . \quad (73)$$

It will be appreciated that the evolution of A_{ab} in the Heisenberg representation is contragradient to the ‘natural’ action of the unitary group $A_{ab} \rightarrow \bar{U}_a^{a'} \bar{U}_b^{b'} A_{a'b'}$ discussed earlier, since the natural action of the unitary group is defined to be the action on A_{ab} that preserves $A_{ab}\xi^a\xi^b$ when ξ^a is evolved in the Schrödinger representation. By the same token, in the Heisenberg representation the action of the time evolution operator on an observable, represented in ‘operator’ form A_a^b (rather than A_{ab}) is given by

$$A_a^b \rightarrow U_a^{a'} A_{a'}^{b'} \bar{U}_{b'}^b , \quad (74)$$

where $U_a^{a'} = \exp[t J_b^{a'} H_a^b]$ and $\bar{U}_a^{a'} = \exp[-t J_b^{a'} H_a^b]$.

Now, finally, we are in a position to address the KMS condition. First we note that for a pair of observables A_{ab} and B_{ab} , their quantum mechanical ‘Dirac’ product $C = AB$ is given by

$$C_{ab} = A_{c(a} \Delta^{cd} B_{b)d} , \quad (75)$$

where $\Delta^{ab} = \frac{1}{2}(g^{ab} - i\Omega^{ab})$. Clearly, C_{ab} is a complex tensor: its real and imaginary parts are given respectively by the Jordan product $\frac{1}{2}(AB + BA)$ and the commutator $\frac{1}{2}i(AB - BA)$. In particular, we have

$$C_{ab} = \frac{1}{2} A_{(a}^c B_{b)}^d g_{cd} - \frac{1}{2} i A_{(a}^c B_{b)}^d \Omega_{cd} . \quad (76)$$

It follows that if ρ_{ab} is a density matrix, the expectation $\rho(AB)$ in the corresponding state is given by $\rho(AB) = \rho_b^a A_c^b \Delta_d^c B_a^d$, or equivalently

$$\rho(AB) = \frac{1}{2} \rho_{ab} A_c^a B^{bc} - \frac{1}{2} i \rho_{ab} A_c^a B_d^b \Omega^{cd} . \quad (77)$$

As we noted earlier, the action of the one-parameter group τ_t on A is given by

$$A_a^b \rightarrow \tau_t A_a^b := \exp \left[t J_a^c H_c^{a'} \right] A_{a'}^{b'} \exp \left[-t J_{b'}^d H_d^b \right] . \quad (78)$$

The KMS complexification of this action, corresponding to replacing t with $t + i\beta$, replaces $t J_a^b$ with $t J_a^b - \beta \delta_a^b$, and we have

$$\tau_{t+i\beta} A_a^b := \exp \left[t J_a^c H_c^{a'} - \beta H_a^{a'} \right] A_{a'}^{b'} \exp \left[-t J_{b'}^d H_d^b + \beta H_{b'}^b \right] . \quad (79)$$

It is then straightforward to verify that in the case of a thermal state, for which

$$\rho_a^b = \frac{\exp[-\beta H_a^b]}{\delta_a^c \exp[-\beta H_c^d]} , \quad (80)$$

the KMS condition $\rho_\beta(\tau_t AB) = \rho_\beta(B \tau_{t+i\beta} A)$ is indeed satisfied.

It might be argued that the complexification $t \rightarrow t + i\beta$ has an artificial character when seen from a real point of view, since it involves an operator transformation of the form $t \delta_a^b \rightarrow t \delta_a^b + \beta J_a^b$. Moreover, equation (79) takes one outside the category of symmetric tensors. Since we would like to argue that the ‘real’ approach to quantum theory acts as the natural bridge between quantum dynamics, on the one hand, and modern statistical theory and hence thermal physics, on the other hand, we are thus led to the negative conclusion that the KMS condition, despite its historical significance in the development of modern thermal physics, should not be regarded as fundamental. This is in keeping with our emphasis on so-called primitive thermal states, which play a significant role even before the consideration of ensemble behaviour.

D. Phase space dynamics and temperature

In the foregoing formulation, we have adopted the view point that, given a system in heat bath with inverse temperature β , we let the equilibrium thermal states evolve quantum mechanically under the influence of the Hamiltonian. We have also considered the case of an ensemble of particles in a thermalised box, whereby while each particle takes a definite energy value, the probability law of the energy for a random element of the ensemble being given by the Boltzmann distribution. In either case, the value of β is taken as an ‘input’ variable that specifies the thermal state. That is to say, we have adopted the standard canonical ensemble of distributions for which temperature is operationally defined.

This point of view can, to some extent, be inverted by shifting gears from the standard canonical description to the microcanonical ensemble. In this case, the essential equivalence of the two approaches is typically recognised only for large systems (see, e.g., [26]). The dynamical formulation we provide here will be useful, in particular, for small systems such as those on a quantum or mesoscopic scale.

First, we regard as given a Hamiltonian function H defined on the quantum mechanical state space CP^n , which is viewed as a real manifold of dimension $2n$. We then foliate the real state manifold with constant energy surfaces, given by $H(x) = E$. The volume $\mathcal{V}(E)$ of such a surface \mathcal{E}_E^{2n-1} , given by

$$\mathcal{V}(E) = \int_{\mathcal{E}_E^{2n-1}} d\sigma^a \frac{\nabla_a H}{\|\text{grad} H\|} , \quad (81)$$

then tells us the number of microscopic states having the energy E . Here, we have written $d\sigma^a = g^{ab} \epsilon_{bc\dots d} dx^c \dots dx^d$ for the natural vector-valued $(2n-1)$ -form on CP^n (viewed as a real manifold). Therefore, using the Boltzmann relation, the entropy $S(E)$ associated with the energy surface \mathcal{E}_E is given by $S(E) = \ln \mathcal{V}(E)$. As a result, the inverse temperature can be calculated from the usual prescription

$$\beta = \frac{dS(E)}{dE} . \quad (82)$$

In this way, we can calculate the temperature of the system directly from the given quantum mechanical dynamics of the system, or equivalently, given the inter-relationship of the energy surfaces.

The construction just noted for obtaining temperature from the underlying phase space dynamics is well known in classical statistical mechanics (see, e.g., [27]). The novelty here, however, is to regard the quantum mechanical state space CP^n as a real $2n$ -dimensional manifold playing the role of the quantum mechanical phase space, hence allowing us to obtain the temperature for quantum mechanical dynamical systems in the case of the microcanonical ensemble. Based upon this formulation, we can also investigate the possibility of various generalisations of the temperature, when the underlying quantum mechanical dynamics is modified. There are a number of distinct generalisations that can be pursued in this context.

First, we can replace the Hamiltonian function defined on the state manifold by a more general observable. In this case, the underlying theory would correspond to the Kibble-type theories of nonlinear quantum mechanics, for which the above prescription would provide the temperature in the case of a nonlinear dynamical system of this sort. Secondly, we can consider replacing the Fubini-Study metric on the state space by a more general metric. Alternatively, a more radical extension is obtained by replacing the state space manifold itself with a general Kähler manifold. In such generalisations the notion of particle states may no longer survive. On the face of it, this might appear to be an undesirable feature. However, our view is that at high energies, as in the instance of particle collisions, the notion of particle states as such may be lost in any case, and recovered only at an asymptotic level, i.e., on the tangent space. For such generalisations the notion of temperature we have outlined above would nonetheless survive, indicating a strong interlink between the geometrical structure of the state manifold and the thermodynamics of associated statistical ensembles.

E. Spin one-half particle

In this section, we study the thermal dynamics of systems having two energy levels. For such systems, it is easy to see that most physical quantities of interest depend only on the energy difference but not on the actual values of the energies. Therefore, for practical purposes any two level system can be viewed as essentially equivalent to a system consisting of a spin one-half particle interacting with an external field.

We study the classical situation first. In this case, we have an Ising spin in a constant magnetic field whose strength is h . The associated real Hilbert space \mathcal{H}^2 is two dimensional,

with orthogonal axes given by the spin-up state u_{\uparrow}^a and the spin-down state u_{\downarrow}^a . If we view \mathcal{H}^2 as an x - y plane, then the infinite temperature ($\beta = 0$) state is the intersection of the segment S_+^1 of the unit circle with the line $y = x$. As the temperature decreases, the equilibrium state moves towards u_{\uparrow}^a or u_{\downarrow}^a depending on the sign of the magnetic field h .

The unnormalised thermal state ψ^a in \mathcal{H}^2 can be obtained by solving the equation $d\psi^a/d\beta = -\frac{1}{2}H_b^a\psi^b$, with the result

$$\psi^a(\beta) = e^{\beta h}u_{\uparrow}^a + e^{-\beta h}u_{\downarrow}^a. \quad (83)$$

This can be normalised by dividing the right hand side by the partition function $Q(\beta) = 2\cosh(\beta h)$. We now project this space down to RP^1 , which is effectively a circle. Since a circle in \mathcal{H}^2 is a double covering of $RP^1 \sim S^1$, the two eigenstates are mapped to opposite points in RP^1 .

Our geometric ‘quantisation’ procedure for this system is as follows. For the given thermal state $\psi^a(\beta)$ we assign a phase factor, and thus obtain a quantum mechanical state space CP^1 , which is viewed as a sphere S^2 . The north and the south poles of the sphere correspond to the two energy eigenstates, and the great circles passing through these two points correspond to unitary equivalent thermal state space trajectories parameterised by the phase factor. The equator of the sphere, in particular, corresponds to the infinite temperature states. This circle is of course a 1-torus, in accordance with the general description of the infinite temperature state manifold given earlier. The Schrödinger dynamics is given by a rigid rotation of the sphere about the axis that passes through the north and the south poles, the two stationary points. This rotation gives rise to a Killing vector field on the sphere, where the angular velocity is given by the strength of the external field h . Thus, the temperature value specifies the latitude on the sphere and the Schrödinger evolution corresponds to a latitudinal circle.

To pursue this in more detail we introduce a complex Hilbert space with coordinates Z^α ($\alpha = 0, 1$) which we regard as homogeneous coordinates for CP^1 . The complex conjugate of Z^α is the ‘plane’ \bar{Z}_α in CP^1 , which in this dimension is simply a point. The point corresponding to \bar{Z}_α is then given by $\bar{Z}^\alpha = \epsilon^{\alpha\beta}\bar{Z}_\beta$, where $\epsilon^{\alpha\beta}$ is the natural symplectic form on the two-dimensional Hilbert space. The relevant formalism in this case is, of course, equivalent to the standard algebra of two-component spinors. In particular, by use of the spinor identity $2X^{[\alpha}Y^{\beta]} = \epsilon^{\alpha\beta}X_\gamma Y^\gamma$, where $X^\alpha\epsilon_{\alpha\beta} = X_\beta$, we obtain

$$ds^2 = \frac{4Z_\alpha dZ^\alpha \bar{Z}_\beta d\bar{Z}^\beta}{(\bar{Z}_\gamma Z^\gamma)^2} \quad (84)$$

for the Fubini-Study metric in this situation, and

$$Z_\gamma dZ^\gamma = iH_{\alpha\beta}Z^\alpha \bar{Z}^\beta dt \quad (85)$$

for the projective Schrödinger equation. The Hamiltonian H_α^β here has a symmetric representation of the form

$$H_{\alpha\beta} = \frac{2hP_{(\alpha}\bar{P}_{\beta)}}{(\bar{P}_\gamma P^\gamma)} \quad (86)$$

where P^α and \bar{P}^α correspond to the stationary points. In particular, we have $H_\beta^\alpha P^\beta = hP^\alpha$ and $H_\beta^\alpha \bar{P}^\beta = -h\bar{P}^\alpha$, which follow if we bear in mind the identity $\bar{P}_\gamma P^\gamma = -P_\gamma \bar{P}^\gamma$.

By insertion of (86) into (85) and then (85) into (84), we deduce that

$$\left(\frac{ds}{dt}\right)^2 = 16h^2 \text{Prob}[Z^\alpha \rightarrow P^\alpha] \text{Prob}[Z^\alpha \rightarrow \bar{P}^\alpha] , \quad (87)$$

where $\text{Prob}[Z^\alpha \rightarrow P^\alpha]$ is the transition probability from Z^α to the north pole P^α , given by the cross ratio

$$\frac{(\epsilon_{\alpha\beta} Z^\alpha \bar{P}^\beta)(\epsilon_{\gamma\delta} \bar{Z}^\gamma P^\delta)}{(\epsilon_{\alpha\beta} Z^\alpha \bar{Z}^\beta)(\epsilon_{\gamma\delta} P^\gamma \bar{P}^\delta)} = \frac{1}{2}(1 + \cos \theta) . \quad (88)$$

Here, θ is the distance from Z^α to P^α , given by the usual angular coordinate measured down from the north pole. The complementary probability $\text{Prob}[Z^\alpha \rightarrow \bar{P}^\alpha]$ is given by $\frac{1}{2}(1 - \cos \theta)$, and it follows that the velocity of the trajectory through the state space is

$$\frac{ds}{dt} = 2h \sin \theta , \quad (89)$$

a special case of the Anandan-Aharonov relation [23] noted above. For the evolutionary trajectory we obtain

$$Z^\alpha = \cos \frac{\theta}{2} e^{i(ht+\phi)} P^\alpha + \sin \frac{\theta}{2} e^{-i(ht+\phi)} \bar{P}^\alpha , \quad (90)$$

where θ and ϕ are the initial coordinates on the sphere for $t = 0$. If appropriate, Z^α can be normalised by setting $P^\alpha \bar{P}_\alpha = 1$. A short calculation then shows that the expectation E of the energy is given by $E = h \cos \theta$, and that the variance of the energy is given by $h^2 \sin^2 \theta$.

Now, we are in a position to examine the primitive thermal trajectories associated with the spin one-half case. In this case the thermal equation is given by

$$\frac{dZ^\alpha}{d\beta} = -\frac{1}{2} \tilde{H}_\beta^\alpha Z^\beta , \quad (91)$$

where $\tilde{H}_\beta^\alpha = H_\beta^\alpha - \delta_\beta^\alpha H_\nu^\mu Z^\nu \bar{Z}_\mu / Z^\gamma \bar{Z}_\gamma$. By use of (86) we then obtain

$$Z^\alpha = \frac{e^{-\frac{1}{2}\beta h + i\phi} P^\alpha + e^{\frac{1}{2}\beta h - i\phi} \bar{P}^\alpha}{(e^{-\beta h} + e^{\beta h})^{1/2}} , \quad (92)$$

where we assume the normalisation $P^\alpha \bar{P}_\alpha = 1$. This shows that at infinite temperature ($\beta = 0$) the state lies on the equator, given by $Z^\alpha = e^{i\phi} P^\alpha + e^{-i\phi} \bar{P}^\alpha$, where ϕ lies in the interval from 0 to 2π . Zero temperature state is obtained by taking the limit $\beta \rightarrow \infty$, and we find that Z^α approaches \bar{P}^α , the south pole, provided $h > 0$. For the expectation of the energy $E = H_\beta^\alpha Z^\beta \bar{Z}_\alpha$ we have

$$E = h \frac{e^{-\beta h} - e^{\beta h}}{e^{-\beta h} + e^{\beta h}} \quad (93)$$

which, as expected, ranges from 0 to $-h$ as β ranges from 0 to ∞ . We note, in particular, that E is independent of the phase angle ϕ , and that the relation $E = h \tanh(\beta h)$ agrees with the result for a classical spin.

For other observables this is not necessarily the case, and we have to consider averaging over the random state \mathbf{Z}^α obtained by replacing ϕ with a random variable Φ , having a uniform distribution over the interval $(0, 2\pi)$. Then for the density matrix $\rho_\beta^\alpha := E[\mathbf{Z}^\alpha \bar{\mathbf{Z}}_\beta]$, where $E[-]$ is the unconditional expectation, we obtain

$$\rho_\beta^\alpha = \frac{e^{-\beta h} P^\alpha \bar{P}_\beta - e^{\beta h} \bar{P}^\alpha P_\beta}{e^{-\beta h} + e^{\beta h}}. \quad (94)$$

The fact that ρ_β^α has trace unity follows from the normalisation condition $Z^\alpha \bar{Z}_\alpha = 1$, and the identity $Z^\alpha \bar{Z}_\alpha = -Z_\alpha \bar{Z}^\alpha$. For the energy expectation $\rho(H) = H_\beta^\alpha \rho_\alpha^\beta$ we then recover (93).

Alternatively, we can consider the phase-space volume approach considered earlier, by assuming a microcanonical distribution for this system. Now, the phase space volume of the energy surface \mathcal{E}_E^1 (a latitudinal circle) is given by $\mathcal{V}(E) = 2\pi \sin \theta$, where θ is the angle measured from the pole of $CP^1 \sim S^2$. Hence, by use of (82), along with the energy expectation $E = h \cos \theta$, we deduce that the value of the system temperature is

$$\beta(E) = \frac{E}{E^2 - h^2}. \quad (95)$$

Since $E \leq 0$ and $E^2 \leq h^2$, the inverse temperature β is positive. Furthermore, we see that $E = 0$ implies $\theta = \pi/2$, the equator of the sphere, which gives infinite temperature ($\beta = 0$), and $E^2 = h^2$ corresponds to $\theta = \pi$, which gives the zero temperature ($\beta = \infty$) state. The equation of state for this system can also be obtained by use of the standard relation $\beta p = \partial S / \partial \mathcal{V}$. In this case, we obtain the equation of state for an ideal gas, i.e., $p\mathcal{V} = \beta^{-1}$. Explicitly, we have

$$p(E) = -\frac{h}{2\pi E} \sqrt{h^2 - E^2}. \quad (96)$$

The pressure is minimised when the spin aligns with the external field, and is maximised at the equator. We note that for positive energies $E > 0$ the temperature takes negative values. If we take $E < 0$ and then flip the direction of the external field h , this situation can be achieved in practice. Although the concept of such a negative temperature is used frequently in the study of Laser phenomena, it is essentially a transient phenomenon [28], and is thus not as such an objective of thermodynamics.

We note, incidentally, that the distinct energy-temperature relationships obtained here in equation (93) for the canonical ensemble and in (95) for the microcanonical ensemble, have qualitatively similar behaviour. Indeed, for a system consisting of a large number of particles these two results are expected to agree in a suitable limit.

V. DISCUSSION

The principal results of this paper are the following. First, we have formulated a projective geometric characterisation for classical probability states. By specialising then to the canonical ensemble of statistical mechanics, we have been able to determine the main features of thermal trajectories, which are expressed in terms of a Hamiltonian gradient

flow. This flow is then shown to be a special case of a projective automorphism on the state space when it is endowed with the natural RP^n metric. It should be clear that the same formalism, and essentially the same results, apply also to the grand canonical and the pressure-temperature distributions.

The quantum mechanical dynamics of equilibrium thermal states can be studied by consideration of the Hopf-type map $RP^{2n+1} \rightarrow CP^n$, which in the present context allows one to regard the quantum mechanical state space as the base space in a fibre manifold which has the structure of an essentially classical thermal state space. The fact that a projective automorphism on a space of constant curvature can be decomposed into two distinct terms suggests the identification of the Killing term with the Schrödinger evolution of the Hamiltonian gradient flow with respect to the symplectic structure, and the other term with thermal evolution of the Hamiltonian gradient flow with respect to the metric—the former gives rise to a linear transformation, while the latter is nonlinear.

There are a number of problems that still remain. First, much of our formulation has been based on the consideration of finite dimensional examples. The study of phase transitions, however, will require a more careful and extensive treatment of the infinite dimensional case. Our analysis of the projective automorphism group, on the other hand, suggests that the infinite dimensional case can also be handled comfortably within the geometric framework. Also, for most of the paper we have adopted the Schrödinger picture, which has perhaps the disadvantage of being inappropriate for relativistic covariance. It would be desirable to reformulate the theory in a covariant manner, in order to study the case of relativistic fields.

Nevertheless, as regards the first problem noted above, the present formulation is sufficiently rich in order to allow us to speculate on a scenario for the spontaneous symmetry breaking of, say, a pure gauge group, in the infinite dimensional situation. In such cases the hypersurface of the parameter space (a curve in the one-parameter case considered here) proliferates into possibly infinite, thermally inequivalent hypersurfaces, corresponding to the multiplicity of the ground state degeneracy, at which the symmetry is spontaneously broken. The hyper-line characterising the proliferation should presumably be called the spinodal line (cf. [29]), along which the Riemann curvature of the parameter space manifold is expected to diverge. Furthermore, a pure thermal state in the ‘high temperature’ region should evolve into a mixed state, obtained by averaging over all possible surfaces, by passing through the geometrical singularity (the spinodal boundary) characterising phase transitions. By a suitable measurement that determines which one of the ground states the system is in, this mixed state will reduce back to a pure state. In a cosmological context this proliferation may correspond, for example, to different θ -vacuums [30]. It is interesting to note that the situation is analogous to the choice of a complex structure [31] for the field theory associated with a curved space-time, as the universe evolves.

In any case, the remarkable advantages of the use of projective space should be stressed. As we have observed, the structure of projective space allows us to identify probabilistic operations with precise geometric relations. One of the problems involved in developing nonlinear (possibly relativistic) generalisations of quantum mechanics concerns their probabilistic interpretation. Formulated in a projective space, such generalisations can be obtained, for example, by replacing the Hamiltonian function by a more general function, or by introducing a more general metric structure. In this way, the assignment of a suitable probability theory can be approached in an appropriate way. In particular, as we have ob-

served, the canonical ensemble of statistical mechanics has an elegant characterisation in projective space—but this is an example of a theory that is highly nonlinear and yet purely probabilistic.

It is also interesting to observe that the nonlinear generalisation of quantum mechanics considered by Kibble [3] and others can be applied to the thermal situation, in the sense that Hamiltonian function defined on the state manifold can be replaced by a general observable. For such generalisations it is not clear what physical interpretation can be assigned. Naively, one might expect that by a suitable choice of an observable the resulting trajectory characterises some kind of nonequilibrium process.

One of the goals of this paper has been to formulate quantum theory at finite temperature in such a way as to allow for the possibility of various natural generalisations. These might include, for example, the stochastic approach to describe measurement theory, or nonlinear relativistic extensions of standard quantum theory, as noted above. These generalisations will be pursued further elsewhere.

Acknowledgement. DCB is grateful to Particle Physics and Astronomy Research Council for financial support.

* Electronic address: d.brody@damtp.cam.ac.uk

† Electronic address: lane@ml.com

1. Haag, R. and Kastler, D., J. Math. Phys. **5**, 848 (1964).
2. Landsmann, N.P., *Aspects of Classical and Quantum Mechanics* (Springer, New York 1998).
3. Kibble, T.W.B., Commun. Math. Phys. **65**, 189 (1979).
4. Gibbons, G.W., J. Geom. Phys. **8**, 147 (1992).
5. Hughston, L.P., in *Twistor Theory*, ed. Huggett, S. (Marcel Dekker, Inc., New York 1995).
6. Brody, D.C. and Hughston, L.P., “Statistical Geometry”, gr-qc/9701051.
7. Brody, D.C. and Hughston, L.P., “Geometry of Thermodynamic States”, quant-ph/9706030.
8. Brody, D.C. and Hughston, L.P., Phys. Rev. Lett. **77**, 2851 (1996).
9. Gibbons, G.W., Class. Quant. Grav. **14**, A155 (1997).
10. Haag, R., Hugenholtz, N.M. and Winnink, M., Commun. math. Phys. **5**, 215 (1967).
11. Xia, D.X., *Measure and Integration Theory on Infinite-Dimensional Spaces* (Academic Press, New York 1972).
12. Mandelbrot, B., Ann. Math. Stat. **33**, 1021 (1962); Phys. Today, (January 1989).
13. Feshbach, H., Phys. Today, (November 1987); Kittel, C., Phys. Today, (May 1988).

14. Ross, S., *Stochastic Process*, (Wiley, New York 1996).
15. Kibble, T.W.B., Commun. Math. Phys. **64**, 239 (1978).
16. Weinberg, S., Phys. Rev. Lett. **62**, 485 (1989).
17. Gisin, N. and Percival, I., Phys. Lett. **A167**, 315 (1992).
18. Hughston, L.P., Proc. Roy. Soc. London **452**, 953 (1996).
19. Hughston, L.P. and Hurd, T.R., Phys. Rep. **100**, 273 (1983).
20. Kobayashi, S. and Nomizu, K. *Foundations of differential geometry*, vols. 1 and 2 (Wiley, New York, 1963 and 1969).
21. Tomonaga, Y., *Riemannian Geometry*, (Kyoritsu Publishing Co., Tokyo 1970).
22. Hughston, L.P. and Tod, K.P., *An Introduction to General Relativity*, (Cambridge University Press, Cambridge 1990).
23. Anandan, J. and Aharonov, Y., Phys. Rev. Lett. **65**, 1697 (1990).
24. Ashtekar, A. and Schilling, T.A., “Geometrical Formulation of Quantum Mechanics”, gr-qc/9706069.
25. Bratteli, O. and Robinson, D.W., *Operator algebras and quantum statistical mechanics*, vols. I and II (Springer, New York 1979,1981).
26. Ruelle, D., *Statistical Mechanics: Rigorous Results* (Addison-Wesley, New York 1989).
27. Thompson, C.J., *Mathematical Statistical Mechanics* (Princeton University Press, Princeton 1972).
28. Kubo, R., *Statistical Mechanics*, (Kyoritsu Publishing Co., Tokyo 1951).
29. Brody, D. and Rivier, N., Phys. Rev. E **51**, 1006 (1995).
30. Sacha, R.G., Phys. Rev. Lett. **78**, 420 (1997).
31. Gibbons, G.W. and Pohle, H.J., Nucl. Phys. **B410**, 117 (1993).